DUKE MATHEMATICAL JOURNAL

VOLUME 6
1940

DUKE UNIVERSITY PRESS
DURHAM, N. C.

CONTENTS

VOLUME 6, 1940

AN INVERSION FORMULA FOR THE LAPLACE INTEGRAL

By R. P. Boas, Jr. and D. V. Widder

Introduction. A function $f(s)$ which is represented by a Laplace-Stieltjes integral

$$f(s) = \int_0^\infty e^{-st}\, d\alpha(t), \tag{1}$$

being analytic in a half-plane $\sigma > \sigma_c$ ($s = \sigma + i\tau$), is uniquely determined by its values in certain parts of that domain. We name four cases:

(a) the values of $f(s)$ and all its derivatives at a single point s_0, $\sigma_0 > \sigma_c$;

(b) the values of $f(s)$ and all its derivatives on the axis of reals in a neighborhood of infinity, $\tau = 0$, $\sigma > \sigma_1$;

(c) the values of $f(s)$ on a vertical line $\sigma = c$, $-\infty < \tau < \infty$;

(d) the values of $f(s)$ on the axis of reals $\tau = 0$, $\sigma_c < \sigma < \infty$.

In any of these cases it should be possible then to determine $\alpha(t)$ uniquely in terms of the stated values of $f(s)$ or its derivatives. The first case has been treated by use of Laguerre polynomials.[1] The second case is handled by the Post-Widder inversion formula[2]

$$\alpha(t) - \alpha(0+) = \lim_{k\to\infty} \int_{k/t}^\infty (-1)^{k+1} f^{(k+1)}(u) \frac{u^k}{k!}\, du.$$

Case (c) is the classical case:

$$\alpha(t) = \lim_{R\to\infty} \frac{1}{2\pi i} \int_{c-iR}^{c+iR} f(s) \frac{e^{st}}{s}\, ds.$$

Case (d) has been treated by Paley and Wiener[3] and by Doetsch.[4] It is the object of the present paper to provide a new inversion formula for this case.

Received December 15, 1939. The first-mentioned author is a National Research Fellow.

[1] D. V. Widder, *An application of Laguerre polynomials*, this Journal, vol. 1(1935), pp. 126–136.

A. G. Domínguez, *Sur les intégrales de Laplace*, Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences, Paris, vol. 205(1937), pp. 1035–1038.

[2] E. L. Post, *Generalized differentiation*, Transactions of the American Mathematical Society, vol. 32(1930), pp. 723–793.

D. V. Widder, *The inversion of the Laplace integral and the related moment problem*, Transactions of the American Mathematical Society, vol. 36(1934), pp. 107–200.

[3] R. E. A. C. Paley and N. Wiener, *Fourier Transforms in the Complex Domain*, New York, 1934, p. 43.

[4] G. Doetsch, *Bedingungen für die Darstellbarkeit einer Funktion als Laplace-Integral und eine Umkehrformel für die Laplace-Transformation*, Mathematische Zeitschrift, vol. 42(1937), pp. 263–286.

Our method consists essentially in reducing (1) to a Stieltjes integral equation and in applying a known solution to that.[5] Suppose that $\alpha(t)$ is an integral, so that (1) becomes

$$f(x) = \int_0^\infty e^{-xt}\varphi(t)\,dt. \tag{2}$$

Then we have formally

$$g(y) = \int_0^\infty e^{-xy}f(x)\,dx = \int_0^\infty e^{-xy}\,dx \int_0^\infty e^{-xt}\varphi(t)\,dt = \int_0^\infty \frac{\varphi(t)}{y+t}\,dt.$$

But it was shown in the paper just quoted that

$$\lim_{k\to\infty} \frac{(-1)^{k-1}}{k!(k-2)!}[t^{2k-1}g^{(k-1)}(t)]^{(k)} = \varphi(t)$$

under certain suitable restrictions on $\varphi(t)$. This clearly leads to the following inversion formula for (2):

$$\varphi(t) = \lim_{k\to\infty} \frac{t^{k-1}}{k!(k-2)!}\int_0^\infty \frac{\partial^k}{\partial u^k}[u^{2k-1}e^{-tu}]f(u)\,du, \tag{3}$$

and it is to be noted that it depends only on a knowledge of the values of $f(x)$ on the axis of reals. Clearly the formula assumes that $\sigma_c \leqq 0$, and certain modifications are necessary when this is not the case.

By use of this inversion formula (3) we obtain new criteria for the solubility of (1). For example, we show that equation (1) has a non-decreasing bounded solution $\alpha(t)$ if and only if $f(s)$ is non-negative, continuous, and bounded on the real axis $0 < \sigma < \infty$, and if

$$\int_0^\infty \frac{\partial^k}{\partial u^k}[u^{2k-1}e^{-tu}]f(u)\,du \geqq 0 \qquad (t > 0; k = 1, 2, \cdots).$$

It is shown that this condition is equivalent to Bernstein's familiar condition, namely, that $f(\sigma)$ is completely monotonic for $0 \leqq \sigma < \infty$. Each set of solubility conditions obtained in terms of the new inversion operation is shown to be equivalent to a known set.

In conclusion we give a new proof of a familiar Tauberian theorem for (1). The method is the classical one devised by Hardy and Littlewood. We believe our proof will be of interest, in spite of the existence of the modern elegant proofs of Karamata and Wiener, since it eliminates one step in the Hardy-Littlewood mechanism, namely, the successive differentiation of the equation (1). Why this shortening of the proof is possible should be clear from the fact that (3) involves no derivatives of $f(s)$.

[5] D. V. Widder, *The Stieltjes transform*, Transactions of the American Mathematical Society, vol. 43(1938), pp. 7–60.

R. P. Boas and D. V. Widder, *The iterated Stieltjes transform*, Transactions of the American Mathematical Society, vol. 45(1939), pp. 1–72.

1. **Inversion of the Laplace integral.** We introduce two definitions.

DEFINITION 1. For $k = 1, 2, \cdots$,

$$Q_k(t, u) = c_k \frac{\partial^k}{\partial u^k}[u^{2k-1}e^{-tu}],$$

$$c_1 = 1, \qquad c_k = \frac{1}{k!\,(k-2)!} \qquad (k = 2, 3, \ldots).$$

DEFINITION 2. For a specified measurable function $f(x)$,

$$A_k(c, t) = e^{ct}t^{k-1}\int_0^\infty Q_k(t, u)f(u + c)\,du \qquad (k = 1, 2, \ldots),$$

if the integral exists (as the limit of a Lebesgue integral).

We quote as a lemma the following known result.

LEMMA 1.1.[6] *If $\alpha(t) \in L(0, R)$ for every positive R, and if*

$$\int_0^\infty \frac{a(t)\,dt}{x + t}$$

converges for some (and hence for all) positive x, then

$$\lim_{k\to\infty} (2k - 1)!\, c_k t^{k-1} \int_0^\infty \frac{v^k a(v)}{(t + v)^{2k}}\,dv = \lim_{k\to\infty} (2k - 1)!\, c_k t^k \int_0^\infty \frac{v^{k-1} a(v)}{(t + v)^{2k}}\,dv$$
$$= a(t)$$

for almost all t in $(0, \infty)$; and, in particular, at every point t where $a(t+)$ and $a(t-)$ exist and $2a(t) = a(t+) + a(t-)$.

THEOREM 1.2. *If*

$$f(x) = \int_0^\infty e^{-xt}\varphi(t)\,dt,$$

the integral converging for $x > x_0$, then for each $c > x_0$

$$\lim_{k\to\infty} A_k(c, t) = \varphi(t)$$

for almost all t in $(0, \infty)$; and, in particular, at every point of continuity of $\varphi(t)$.

We must first show that $A_k(c, t)$ is well defined for $t > 0$. We write

$$\Phi_r(t) = \int_0^t e^{-ru}\varphi(u)\,du.$$

Then

$$f(x) = \int_0^\infty e^{-(x-r)t}\,d\Phi_r(t), \qquad x > x_0. \tag{1.1}$$

If $r > x_0$, $\Phi_r(\infty) = f(r)$, so that for each $r > x_0$

[6] D. V. Widder, *The Stieltjes transform*, loc. cit., pp. 19, 21.

$$\Phi_r(t) = O(1) \qquad (t \to \infty). \tag{1.2}$$

Hence we may integrate by parts in (1.1), obtaining

$$\begin{aligned} f(x) &= (x - r) \int_0^\infty e^{-(x-r)t} \Phi_r(t)\, dt \qquad (r > x_0) \\ &= O(1) \qquad (x \to \infty). \end{aligned}$$

This result, together with the continuity of $f(u + c)$ for $u \geqq 0$ (if $c > x_0$), is enough to show that the integral in Definition 2 exists if $c > x_0$, $t > 0$, $k = 1, 2, \cdots$.

By (1.2), $\Phi_c(t) = O(1)$ as $t \to \infty$, if $c > x_0$. We then have, by an integration by parts,

$$\begin{aligned} A_k(c, t) &= c_k t^{k-1} e^{ct} \int_0^\infty \frac{\partial^k}{\partial u^k} [u^{2k-1} e^{-tu}] f(u + c)\, du \\ &= (-1)^k c_k t^{k-1} e^{ct} \int_0^\infty u^{2k-1} e^{-tu} f^{(k)}(u + c)\, du \\ &= c_k t^{k-1} e^{ct} \int_0^\infty u^{2k-1} e^{-tu}\, du \int_0^\infty v^k e^{-(u+c)v} \varphi(v)\, dv \\ &= c_k t^{k-1} e^{ct} \int_0^\infty u^{2k-1} e^{-tu}\, du \int_0^\infty v^k e^{-uv}\, d\Phi_c(v) \\ &= c_k t^{k-1} e^{ct} \int_0^\infty u^{2k-1} e^{-tu}\, du \int_0^\infty \Phi_c(v) e^{-uv} v^{k-1}(uv - k)\, dv. \end{aligned}$$

The last iterated integral is dominated by the convergent iterated integral

$$\int_0^\infty u^{2k-1} e^{-tu}\, du \int_0^\infty e^{-uv} v^{k-1}(uv + k)\, dv = 2(k!) \int_0^\infty u^{k-1} e^{-tu}\, du.$$

Hence the order of integration can be changed, and we obtain

$$\begin{aligned} A_k(c, t) &= c_k t^{k-1} e^{ct} \int_0^\infty \Phi_c(v) v^{k-1}\, dv \int_0^\infty u^{2k-1} e^{-(t+v)u}(uv - k)\, du \\ &= -(2k - 1)! c_k t^{k-1} e^{ct} \int_0^\infty \frac{\Phi_c(v) v^{k-1}}{(t + v)^{2k}} \left(k - \frac{2kv}{t + v}\right) dv \\ &= (2k - 1)! c_k t^{k-1} e^{ct} \int_0^\infty \frac{v^k}{(t + v)^{2k}}\, e^{-cv} \varphi(v)\, dv. \end{aligned}$$

Now, for $x > 0$ and $c > x_0$,

$$\int_0^R \frac{\varphi(v) e^{-cv}\, dv}{x + v} = \int_0^R \frac{d\Phi_c(v)}{x + v} = \frac{\Phi_c(R)}{x + R} + \int_0^R \frac{\Phi_c(v)\, dv}{(x + v)^2},$$

which is seen, by (1.2), to approach a limit as $R \to \infty$. Consequently Lemma 1.1 applies to the function $a(v) = e^{-cv}\varphi(v)$, and the conclusion of the theorem follows.

To invert the Laplace-Stieltjes integral

$$f(x) = \int_0^\infty e^{-xt}\, d\alpha(t), \qquad \alpha(0) = 0,$$

we may put it into the form

$$f(x) = x\int_0^\infty \alpha(t)e^{-xt}\, dt,$$

and apply Theorem 1.3 to $f(x)/x$. A more elegant inversion formula is furnished by the following theorem.

THEOREM 1.3. *If $\alpha(t)$ is a normalized function of bounded variation in $0 \leqq t \leqq R$ for every positive R, and if*

$$f(x) = \int_0^\infty e^{-xt}\, d\alpha(t),$$

the integral converging for $x > x_0$, then for each $c > x_0$

$$\lim_{k\to\infty} \int_0^t A_k(c, u)\, du = \alpha(t) - \alpha(0+) \qquad (t > 0).$$

Define

$$\Phi_c(t) = \int_0^t e^{-cu}\, d\alpha(u).$$

Then

$$\begin{aligned} A_k(c, t) &= c_k t^{k-1} e^{ct} \int_0^\infty u^{2k-1} e^{-tu}\, du \int_0^\infty v^k e^{-uv}\, d\Phi_c(v) \\ &= c_k t^{k-1} e^{ct} \int_0^\infty v^k\, d\Phi_c(v) \int_0^\infty e^{-u(v+t)} u^{2k-1}\, du. \end{aligned}$$

The change in the order of integration may be made as in the proof of Theorem 1.2 by use of integration by parts. It follows that

$$A_k(c, t) = (2k-1)!\, c_k t^{k-1} e^{ct} \int_0^\infty \frac{v^k}{(v+1)^{2k}}\, e^{-cv}\, d\alpha(v).$$

Integrating by parts, we have

$$A_k(c, t) = -(2k-1)!\, c_k \int_0^\infty \Phi_c(v)\, \frac{\partial}{\partial v}\left[\frac{t^{k-1} v^k e^{ct}}{(v+t)^{2k}}\right] dv.$$

By direct computation we have

$$\frac{\partial}{\partial v}\left[\frac{t^{k-1} v^k e^{ct}}{(v+t)^{2k}}\right] = -\frac{\partial}{\partial t}\left[\frac{t^k v^{k-1} e^{ct}}{(v+t)^{2k}}\right] + ct\, \frac{e^{ct} v^{k-1} t^{k-1}}{(v+t)^{2k}}.$$

Moreover, the integral

$$\int_0^\infty \frac{\partial}{\partial t}\left[\frac{t^k v^{k-1} e^{ct}}{(v+t)^{2k}}\right] \Phi_c(v)\, dv$$

is easily seen to converge uniformly in the interval $\epsilon \leqq t \leqq R$ for any positive numbers ϵ and R. One has only to make use of the relation

$$\Phi_c(v) = O(1) \qquad (v \to \infty)$$

to prove this fact. Consequently we may integrate under the integral sign and obtain

$$\begin{aligned}\int_\epsilon^R A_k(c, t)\,dt &= (2k-1)!c_k \int_0^\infty \frac{R^k v^{k-1} e^{cR}}{(R+v)^{2k}} \Phi_c(v)\,dv \\ &\quad - (2k-1)!c_k \int_0^\infty \frac{\epsilon^k v^{k-1} e^{c\epsilon}}{(\epsilon+v)^{2k}} \Phi_c(v)\,dv \\ &\quad + (2k-1)!c_k \int_\epsilon^R dt \int_0^\infty \frac{v^{k-1} t^{k-1} e^{ct}}{(v+t)^{2k}} ct\Phi_c(v)\,dv \\ &= I_1 + I_2 + I_3.\end{aligned}$$

By a familiar Abelian argument[7]

$$\lim_{\epsilon \to 0+} I_2 = -\frac{k-1}{k} \Phi_c(0+) = -\frac{k-1}{k} \alpha(0+).$$

Hence

$$\int_{0+}^R A_k(c, t)\,dt = I_1 - \frac{k-1}{k} \alpha(0) + \lim_{\epsilon \to 0+} I_3.$$

But by Lemma 1.1

$$\lim_{k \to \infty} (2k-1)!c_k \int_0^\infty \frac{v^{k-1} t^{k-1}}{(v+t)^{2k}} ct\Phi_c(v) e^{ct}\,dv = 0. \tag{1.3}$$

Since there exists a constant M such that

$$|\Phi_c(v)| < M \qquad (0 \leqq v < \infty),$$

we have

$$\left|(2k-1)!c_k \int_0^\infty \frac{v^{k-1} t^{k-1}}{(v+t)^{2k}} ct\Phi_c(v) e^{ct}\,dv\right| \leqq Mc \frac{2k-1}{k} e^{ct} < 2Mce^{ct} \quad (0 \leqq t < \infty).$$

Hence the limit (1.3) is attained boundedly in $(0, R)$ so that by Lebesgue's limit theorem

$$\lim_{k \to \infty} \int_{0+}^R dt \int_0^\infty \frac{v^{k-1} t^{k-1}}{(v+t)^{2k}} cte^{ct} \Phi_c(v)\,dv = 0.$$

Another application of Lemma 1.1 gives

$$\lim_{k \to \infty} I_1 = \alpha(R).$$

[7] See D. V. Widder, *The Stieltjes transform*, loc. cit., p. 18, Lemma 3.12.

Hence

$$\lim_{k\to\infty} \int_{0+}^{R} A_k(c, t)\, dt = \alpha(R) - \alpha(0+).$$

To show that $\int_{0+}^{R}$ can be replaced by $\int_{0}^{R}$, we need only show that $\int_{0}^{1}$ exists. We have

$$|A_k(c, t)| \leqq (2k-1)!c_k t^{k-1} e^{ct}\left\{\int_0^1 \frac{v^k e^{-cv}}{(v+t)^{2k}}\, |d\alpha(v)| + \left|\int_1^\infty \frac{v^k e^{-cv}}{(v+t)^{2k}}\, d\alpha(v)\right|\right\}$$
$$= B(t) + C(t).$$

Then

$$\frac{1}{(2k-1)!c_k}\int_0^1 B(t)\, dt = \int_0^1 t^{k-1}e^{ct}\, dt \int_0^1 \frac{v^k e^{-cv}}{(v+t)^{2k}}\, |d\alpha(v)|$$
$$= \int_0^1 v^k e^{-cv}\, |d\alpha(v)| \int_0^1 \frac{t^{k-1}\, dt}{(v+t)^{2k}}$$
$$\leqq \frac{(k-1)!(k-1)!}{(2k-1)!}\int_0^1 e^{-cv}\, |d\alpha(v)| < \infty,$$

since $\alpha(v)$ has bounded variation on $(0, 1)$; the change in the order of integration is justified by the Fubini theorem for Stieltjes integrals.[8]

To discuss $C(t)$, we write

$$\frac{C(t)}{(2k-1)!c_k} = t^{k-1}e^{ct}\left|\int_1^\infty \frac{v^k}{(t+v)^{2k}}\, d\Phi_c(v)\right| \leqq \frac{t^{k-1}e^{ct}\, |\Phi_c(1)|}{(1+t)^{2k}}$$
$$+ t^{k-1}e^{ct}\left|\int_1^\infty \Phi_c(v)\, \frac{\partial}{\partial v}\left[\frac{v^k}{(v+t)^{2k}}\right] dv\right|.$$

Since $\Phi_c(\infty)$ exists, there is a constant M such that $|\Phi_c(v)| \leqq M$ for $1 \leqq v < \infty$. Since $v(v+t)^{-2}$ decreases if $v \geqq t$, and since $t \leqq 1$, we therefore have

$$\frac{C(t)}{(2k-1)!c_k} \leqq \frac{Mt^{k-1}e^{ct}}{(1+t)^{2k}} - Mt^{k-1}e^{ct}\int_1^\infty \frac{\partial}{\partial v}\left[\frac{v^k}{(v+t)^{2k}}\right] dv = \frac{2Mt^{k-1}e^{ct}}{(1+t)^{2k}}.$$

Consequently $\int_0^1 C(t)\, dt$ exists.

2. **A general limit theorem.** We now establish two theorems which will enable us to give new proofs of several criteria for the representation of functions by Laplace integrals.

THEOREM 2.1. *If $f(x)$ is a measurable function such that for some numbers $\delta > 0, \Delta > 0$,*

[8] S. Saks, *Theory of the Integral*, Warsaw, 1937, p. 87.

$$f(x) = O\left\{\frac{1}{(x-c)^{\Delta}}\right\}, \qquad x \to c+, \tag{2.1}$$

and

$$f(x) = O(x^{-\delta}), \qquad x \to \infty, \tag{2.2}$$

then

$$\lim_{k\to\infty} \int_0^\infty e^{-xt} A_k(c, t)\, dt = f(x)$$

for almost all x in (c, ∞), and in particular at every point of continuity $x_0 > c$ of $f(x)$.

More generally, we could replace (at the expense of additional complication in the proof) condition (2.2) by

$$f(x) \in L^p(c + 1, \infty)$$

for some p, $1 \leqq p < \infty$; and (2.1) by a similar condition.

We see at once from Definition 2 that $A_k(c, t)$ is well defined for $t > 0$, $k = 1, 2, \cdots$ if (2.1) and (2.2) are satisfied. We now consider

$$I_k(x) = \int_0^\infty e^{-xt} A_k(c, t)\, dt = \int_0^\infty e^{(c-x)t} t^{k-1}\, dt \int_0^\infty Q_k(t, u) f(u + c)\, du, \quad x > c. \tag{2.3}$$

We wish to show that the iterated integral is absolutely convergent. If we carry out the indicated differentiations in the definition of $Q_k(t, u)$, we see that it is enough to establish the convergence of

$$\int_0^\infty |f(u + c)|\, du \int_0^\infty e^{(c-x-u)t} (tu)^r\, dt \qquad (r = k - 1, k, \cdots, 2k - 1);$$

or of

$$\int_0^\infty \frac{u^r |f(u + c)|\, du}{(c - x - u)^{r+1}} \qquad (r = k - 1, k, \cdots, 2k - 1).$$

That these integrals converge is a consequence of (2.1) and (2.2) if $k > 1 + \Delta$.

We may now change the order of integration in (2.3), and write $I_k(x)$ in the form

$$\begin{aligned} I_k(x) &= c_k \int_0^\infty f(u + c)\, du \int_0^\infty e^{(c-x)t} t^{k-1} \frac{\partial^k}{\partial u^k} [u^{2k-1} e^{-tu}]\, dt \\ &= c_k \int_0^\infty f(u + c) \frac{\partial^k}{\partial u^k}\left\{u^{2k-1} \int_0^\infty e^{(c-x-u)t} t^{k-1}\, dt\right\} du \\ &= (k - 1)!\, c_k \int_0^\infty f(u + c) \frac{\partial^k}{\partial u^k}\left\{\frac{u^{2k-1}}{(u + x - c)^k}\right\} du. \end{aligned}$$

To evaluate $\dfrac{\partial^k}{\partial u^k}\{\cdots\}$, we recall that[9]

$$\frac{d^k}{du^k}\{u^{2k-1}F^{(k-1)}(u)\} = u^{k-1}\frac{d^{2k-1}}{du^{2k-1}}\{u^kF(u)\}.$$

Applying this identity to $F(u) = (x + u - c)^{-k}$, we obtain

$$(-1)^k(k-1)!\frac{\partial^k}{\partial u^k}\left\{\frac{u^{2k-1}}{(x+u-c)^k}\right\} = u^{k-1}\frac{\partial^{2k-1}}{\partial u^{2k-1}}\left\{\frac{[u+x-c+(c-x)]^k}{u+x-c}\right\}$$
$$= \frac{(-1)^{k-1}(2k-1)!\,u^{k-1}(x-c)^k}{(u+x-c)^{2k}},$$

since

$$\frac{\{u+x-c+(c-x)\}^k}{u+x-c} = \frac{(c-x)^k}{u+x-c} + \sum_{j=1}^{k}\binom{k}{j}(u+x-c)^{j-1}(c-x)^{k-j}.$$

Hence we have

$$I_k(x) = (2k-1)!\,c_k(x-c)^k\int_0^\infty \frac{f(u+c)u^{k-1}}{(x+u-c)^{2k}}\,du$$
$$\to f(x)$$

for almost all x in (c, ∞), by Lemma 1.1 with $a(u) = f(u+c)$, $t = x - c$.

THEOREM 2.2. *If $f(x)$ is measurable on (c, ∞), essentially bounded on $(c+1, \infty)$, and*

$$f(x) = O\{(x-c)^{-\Delta}\}, \qquad x \to c+,$$

for some $\Delta > 0$, then for $h > 0$ and $k > \Delta + 1$,

$$\int_0^\infty e^{-xt}(1-e^{-ht})A_k(c,t)\,dt$$
$$(2.4)\qquad = (2k-1)!\,c_k\int_0^\infty f(u+c)u^{k-1}\left\{\frac{(x-c)^k}{(x+u-c)^{2k}} - \frac{(x+h-c)^k}{(x+h+u-c)^{2k}}\right\}du,$$

and consequently

$$(2.5)\qquad \lim_{k\to\infty}\int_0^\infty e^{-xt}(1-e^{-ht})A_k(c,t)\,dt = f(x) - f(x+h)$$

for almost all x on (c, ∞), and in particular at every point of continuity $x_0 > c$ of $f(x)$.

The hypotheses are those of Theorem 2.1, but with $\delta = 0$; the proof is similar to that of Theorem 2.1. We consider

$$I_k(x) = \int_0^\infty e^{-xt}(1-e^{-ht})A_k(c,t)\,dt$$
$$(2.6)\qquad = \int_0^\infty e^{(c-x)t}(1-e^{-ht})t^{k-1}\,dt\int_0^\infty Q_k(t,u)f(u+c)\,du, \qquad x > c.$$

[9] D. V. Widder, *The Stieltjes transform*, loc. cit., p. 17.

The iterated integral is dominated by a linear combination of the integrals

$$\int_0^\infty |f(u+c)|\,du \int_0^\infty e^{(c-x-u)t}(1-e^{-ht})(tu)^r\,dt \qquad (r = k-1, k, \cdots, 2k-1),$$

which differ only by numerical factors from

$$\int_0^\infty |f(u+c)|\,u^r\left\{\frac{1}{(x+u-c)^{r+1}} - \frac{1}{(x+u+h-c)^{r+1}}\right\}du.$$

Since the factor in the brace is $O(u^{-r-2})$ as $u \to \infty$, these integrals converge, and we may change the order of integration in (2.6). We then obtain (2.4) and (2.5) as in the proof of Theorem 2.1.

The following theorem is an immediate corollary of Theorem 2.2.

THEOREM 2.3. *Under the hypotheses of Theorem* 2.2,

$$\lim_{k\to\infty} \int_0^\infty \Delta_k^n[e^{-xt}]A_k(c, t)\,dt = \Delta_h^n[f(x)]$$

for almost all x on (c, ∞), if $n \geqq 1$ and $h > 0$. (The differences are taken with respect to x.)

In fact, Theorem 2.2 is the case $n = 1$, and the differences of higher order are linear combinations of those of order one.

3. **A lemma.** We need the following lemma to establish the necessity of our criteria for the representation of functions by Laplace integrals.

LEMMA 3.1. *If*

$$f(x) = \int_0^\infty e^{-xt}\,d\alpha(t), \qquad x > 0, \tag{3.1}$$

with $\alpha(0) = 0$, and

$$V(t) = \int_0^t |d\alpha(u)| = O(t^K), \qquad t \to \infty, \tag{3.2}$$

for some $K \geqq 0$, then for $t > 0$, and $k > K$,

$$A_k(0, t) = (2k-1)!\,c_k t^{k-1} \int_0^\infty \frac{v^k}{(t+v)^{2k}}\,d\alpha(v). \tag{3.3}$$

Since for $k = 0, 1, 2, \cdots$

$$\left|\int_0^\infty t^k e^{-xt}\,d\alpha(t)\right| = \left|\int_0^\infty \alpha(t)t^{k-1}(k - xt)e^{-xt}\,dt\right|$$
$$\leqq \int_0^\infty |\alpha(t)|\,t^{k-1}(k + xt)e^{-xt}\,dt,$$

we see that

$$f^{(k)}(x) = O(x^{-K-k}), \qquad x \to 0+. \tag{3.4}$$

From the formula

$$A_k(0, t) = c_k t^{k-1} \int_0^\infty f(u) \frac{\partial^k}{\partial u^k} [u^{2k-1} e^{-tu}] \, du$$

we then see that $A_k(0, t)$ is well defined for $t > 0$, $k > K$; and we may integrate by parts, the integrated terms vanishing. Thus

$$\begin{aligned} A_k(0, t) &= (-1)^k c_k t^{k-1} \int_0^\infty u^{2k-1} e^{-tu} f^{(k)}(u) \, du \\ &= c_k t^{k-1} \int_0^\infty u^{2k-1} e^{-tu} \, du \int_0^\infty s^k e^{-us} \, d\alpha(s) \\ &= c_k t^{k-1} \int_0^\infty s^k \, d\alpha(s) \int_0^\infty u^{2k-1} e^{-u(t+s)} \, du \\ &= (2k-1)! \, c_k t^{k-1} \int_0^\infty \frac{s^k \, d\alpha(s)}{(t+s)^{2k}}. \end{aligned}$$

The change in the order of integration is justified if the last integral is absolutely convergent. But for $k > K$, by (3.2),

$$\int_0^\infty \frac{s^k \, dV(s)}{(t+s)^{2k}} = k \int_0^\infty V(s) \frac{s^{k-1}}{(t+s)^{2k}} \left(\frac{2s}{t+s} - 1 \right) ds,$$

which converges. This completes the proof.

4. Representation theorems: $\varphi(t)$ bounded.

THEOREM 4.1. *A set of necessary and sufficient conditions for $f(x)$ to have the representation*

$$f(x) = \int_0^\infty e^{-xt} \varphi(t) \, dt, \qquad x > 0, \tag{4.1}$$

with[10]

$$\sup_{0<t<\infty}{}^{\circ} \, | \varphi(t) | \leqq M \tag{4.2}$$

is

(i) *$f(x)$ continuous in* $0 < x < \infty$;

(ii) $f(x) = O\left(\frac{1}{x}\right) \qquad (x \to \infty, x \to 0+)$;

(iii) *for* $k = 2, 3, \cdots$,

$$| A_k(0, t) | \leqq M, \qquad 0 < t < \infty. \tag{4.3}$$

The necessity of condition (ii) is clear, since

$$|f(x)| \leqq M \int_0^\infty e^{-xt} \, dt = \frac{M}{x}.$$

[10] The notation sup° is the "vrai max" of S. Banach, *Théorie des Opérations Linéaires*, Warszawa, 1932, p. 227.

To establish the necessity of (4.3), we use Lemma 3.1, whose hypotheses are obviously satisfied with $K = 1$. We obtain

$$\begin{aligned} | A_k(0, t) | &\leqq M(2k-1)!\,c_k t^{k-1} \int_0^\infty \frac{v^k}{(t+v)^{2k}}\,dv \\ &= M \qquad\qquad (k = 2, 3, \cdots). \end{aligned}$$

To establish the sufficiency of the conditions, we apply Theorem 2.1, whose hypotheses are clearly satisfied. We have

$$f(x) = \lim_{k\to\infty} \int_0^\infty e^{-xt} A_k(0, t)\,dt, \qquad 0 < x < \infty.$$

Since the functions $A_k(0, t)$ are uniformly bounded, there is an essentially bounded function $\varphi(t)$ such that

$$\lim_{k\to\infty} \int_0^\infty g(t) A_k(0, t)\,dt = \int_0^\infty g(t)\varphi(t)\,dt$$

for every $g(t) \in L(0, \infty)$,[11] and in particular for $g(t) = e^{-xt}$. This establishes (4.1); (4.2) is a consequence of (4.3) and Theorem 1.2.

Theorem 4.1 leads to a new proof of the sufficiency part of the following known theorem.

THEOREM 4.2. *A necessary and sufficient condition for $f(x)$ to have the representation* (4.1) *with $\varphi(t)$ satisfying* (4.2) *that $f(x)$ be of class C^∞ in $0 < x < \infty$, and*

$$| f^{(k)}(x) |\, x^{k+1} \leqq k!M \qquad (x > 0;\ k = 0, 1, 2, \cdots). \tag{4.4}$$

If (4.4) is satisfied, condition (ii) of Theorem 4.1 is evident. To establish (4.3), we have

$$A_k(0, t) = c_k t^{k-1} \int_0^\infty f(u) \frac{\partial^k}{\partial u^k} [u^{2k-1} e^{-tu}]\,du, \qquad k \geqq 2. \tag{4.5}$$

If we integrate k times by parts, the integrated terms vanish, by (4.4), and

$$A_k(0, t) = -c_k(-t)^{k-1} \int_0^\infty u^{2k-1} e^{-tu} f^{(k)}(u)\,du,$$

$$\begin{aligned} | A_k(0, t) | &\leqq Mk!\,c_k t^{k-1} \int_0^\infty u^{k-2} e^{-tu}\,du \\ &= M. \end{aligned}$$

The sufficiency part of Theorem 4.2 now follows from Theorem 4.1.

5. Representation theorems: $\varphi(t) \in L^p$, $p > 1$.

THEOREM 5.1. *A set of necessary and sufficient conditions for $f(x)$ to have the representation*

[11] S. Banach, loc. cit., p. 130.

$$f(x) = \int_0^\infty e^{-xt}\varphi(t)\,dt, \qquad x > 0, \tag{5.1}$$

with

$$\int_0^\infty |\varphi(t)|^p\,dt \leqq M^p \qquad (p > 1) \tag{5.2}$$

is

(i) $f(x)$ *continuous in* $0 < x < \infty$;

(ii) *as* $x \to \infty$ *and as* $x \to 0+$

$$f(x) = O(x^{(1-p)/p}); \tag{5.3}$$

(iii) *for* $k = 1, 2, \cdots$

$$\int_0^\infty |A_k(0, t)|^p\,dt \leqq M^p. \tag{5.4}$$

The necessity of (5.3) follows from

$$|f(x)|^p \leqq \int_0^\infty |\varphi(t)|^p\,dt\left\{\int_0^\infty e^{-p'xt}\,dt\right\}^{p-1}$$

$$= \left(\frac{1}{p'x}\right)^{p-1}\int_0^\infty |\varphi(t)|^p\,dt, \qquad p' = \frac{p}{p-1}.$$

To show that (5.4) is necessary, we apply Lemma 3.1, whose hypotheses are satisfied with $K < 1$, since

$$\int_0^t |\varphi(u)|\,du \leqq \left(\int_0^t |\varphi(u)|^p\,du\right)^{1/p}\left(\int_0^t du\right)^{1/p'}$$

$$\leqq Mt^{1/p'}.$$

We have

$$\left\{\int_0^\infty |A_k(0, t)|^p\,dt\right\}^{1/p} \leqq (2k-1)!\,c_k\left\{\int_0^\infty\left(\frac{v^k t^{k-1}}{(t+v)^{2k}}|\varphi(v)|\,dv\right)^p dt\right\}^{1/p}$$

$$\leqq \frac{\Gamma(k+1/p')\Gamma(k-1/p')}{k!(k-2)!}\left(\int_0^\infty |\varphi(v)|^p\,dv\right)^{1/p}$$

by an inequality of Hardy, Littlewood, and Pólya.[12] In fact,

$$K(v, t) = \frac{v^k t^{k-1}}{(t+v)^{2k}}$$

is homogeneous of degree -1, and

$$\int_0^\infty K(v, 1)v^{-1/p}\,dv = \int_0^\infty K(1, t)t^{-1/p'}\,dt = \frac{\Gamma(k+1/p')\Gamma(k-1/p')}{\Gamma(2k)}.$$

[12] G. H. Hardy, J. E. Littlewood, G. Pólya, *Inequalities*, Cambridge, 1934, p. 229.

Finally,

$$\frac{\Gamma(k+1/p')\Gamma(k-1/p')}{k!(k-2)!} = \frac{\Gamma(k+1/p')\Gamma(k-1+1/p)}{k!(k-2)!}$$

$$= \frac{(k-1+1/p')(k-2+1/p')\cdots(1+1/p')\Gamma(1+1/p')}{k(k-1)\cdots 2\cdot 1}$$

$$\cdot\frac{(k-2+1/p)\cdots(1+1/p)\Gamma(1+1/p)}{(k-2)(k-3)\cdots 2\cdot 1} < 1,$$

since $\Gamma(x) < 1$ for $1 < x < 2$.

The proof of the sufficiency proceeds just as in Theorem 4.1; we use the "weak compactness" of the space L^p instead of that of the space B of essentially bounded functions.

From Theorem 5.1 we obtain a new proof of the sufficiency part of the following known theorem.

THEOREM 5.2. *A set of necessary and sufficient conditions for $f(x)$ to have the representation* (5.1) *with $\varphi(t)$ satisfying* (5.2) *is*

(i) $f(x) \in C^\infty$, $0 < x < \infty$;

(ii) $f(\infty) = 0$;

(iii) *for* $k = 1, 2, \cdots,$

$$\frac{k}{(k!)^p}\int_0^\infty |f^{(k)}(u)|^p u^{(k+1)p-2}\,du \leqq M^p. \tag{5.5}$$

For $k = 1, 2, \cdots$, we have, using Hölder's inequality and (5.5),

$$\int_x^\infty |f^{(k)}(t)|t^{k-1}\,dt \leqq \left\{\int_x^\infty |f^{(k)}(t)|^p t^{(k+1)p-2}\,dt\right\}^{1/p}\left\{\int_x^\infty \frac{dt}{t^2}\right\}^{1/p'}$$

$$\leqq Mk!\,k^{-1/p}x^{-1/p'}.$$

This inequality shows that the integrals

$$\int_x^\infty f^{(k)}(t)t^{k-1}\,dt \qquad (k = 1, 2, \ldots)$$

all exist. By an integration by parts we then have

$$\left|\lim_{x\to\infty} x^{k-1}f^{(k-1)}(x) - x^{k-1}f^{(k-1)}(x) - (k-1)\int_x^\infty f^{(k-1)}(t)t^{k-2}\,dt\right|$$

$$\leqq Mk!\,k^{-1/p}x^{-1/p'},$$

$$\begin{aligned}&|\lim_{x\to\infty} x^{k-1}f^{(k-1)}(x) - x^{k-1}f^{(k-1)}(x)|\\ &\qquad\leqq M(k-1)!\,\{k^{1/p'} + (k-1)^{1/p'}\}x^{-1/p'},\end{aligned} \tag{5.6}$$

for $k = 2, 3, \cdots$; and

$$|f(\infty) - f(x)| \leqq Mx^{-1/p'}.$$

By hypothesis (ii) of the theorem, $f(\infty) = 0$, so that

$$f(x) = O(x^{-1/p'}), \qquad x \to \infty.$$

From (5.6) we deduce first that

$$f^{(k)}(x) = O(x^{-k}), \qquad x \to \infty \qquad (k = 1, 2, \cdots); \tag{5.7}$$

then, by an "O-theorem" of Hardy and Littlewood,[13] that

$$f^{(k)}(x) = o(x^{-k}), \qquad x \to \infty \qquad (k = 0, 1, 2, \cdots);$$

and finally, since the limit on the left of (5.6) vanishes, that

$$f^{(k)}(x) = O(x^{-k-1/p'}) = O(x^{-k+(1-p)/p}), \qquad x \to \infty \qquad (k = 0, 1, 2, \cdots). \tag{5.8}$$

From (5.6) we now have also

$$f^{(k)}(x) = O(x^{-k+(1-p)/p}), \qquad x \to 0+ \qquad (k = 0, 1, 2, \cdots). \tag{5.9}$$

The inequalities (5.8) and (5.9) allow us to integrate by parts in the relation

$$A_k(0, t) = c_k t^{k-1} \int_0^\infty f(u) \frac{\partial^k}{\partial u^k} [u^{2k-1} e^{-tu}]\, du, \qquad k \geqq 1,$$

the integrated terms vanishing. Then

$$\begin{aligned} | A_k(0, t) | &\leqq c_k t^{k-1} \int_0^\infty u^{2k-1} e^{-tu} \, | f^{(k)}(u) | \, du, \\ \int_0^\infty | A_k(0, t) |^p \, dt &\leqq c_k^p \int_0^\infty t^{pk-p}\, dt \left\{ \int_0^\infty u^{k-2+2/p} e^{-tu} F(u)\, du \right\}^p, \end{aligned} \tag{5.10}$$

where $F(u) = | f^{(k)}(u) | \, u^{k+1-2/p}$, and

$$\int_0^\infty \{F(u)\}^p \, du \leqq \frac{M^p (k!)^p}{k}, \tag{5.11}$$

since (5.5) is satisfied.

With the change of variable $t = 1/s$, (5.10) can be written

$$\int_0^\infty | A_k(0, t) |^p \, dt \leqq c_k^p \int_0^\infty ds \left\{ \int_0^\infty F(u) K(u, s)\, du \right\}^p, \tag{5.12}$$

where

$$K(u, s) = e^{-u/s} u^{k-2+2/p} s^{1-k-2/p}.$$

$K(u, s)$ is homogeneous of order -1, and a simple calculation shows that

$$\int_0^\infty K(u, 1) u^{-1/p}\, du = \int_0^\infty K(1, s) s^{-1/p'}\, ds = \Gamma(k - 1 + 1/p).$$

An inequality of Hardy, Littlewood, and Pólya[14] is now applicable, and, with (5.11), shows that

[13] G. H. Hardy and J. E. Littlewood, *Contributions to the arithmetic theory of series*, Proceedings of the London Mathematical Society, (2), vol. 11 (1912), pp. 411–478; 426. We do not really need to appeal to this theorem since (5.7) would be enough for our purposes.

$$\begin{aligned}\int_0^\infty |A_k(0, t)|^p dt &\leqq c_k^p \{\Gamma(k - 1 + 1/p\}^p \int_0^\infty \{F(u)\}^p du \\ &\leqq \frac{1}{k} \{Mc_k k!\Gamma(k - 1 + 1/p)\}^p \\ &\leqq \frac{M^p}{k} \left\{\frac{\Gamma(k - 1 + 1/p)}{\Gamma(k - 1)}\right\}^p \\ &\sim M^p, \qquad k \to \infty.\end{aligned}$$

The last statement is easily verified by using Stirling's formula.

We therefore have (5.4), with M replaced by some M'. Since (5.3) is the case $k = 0$ of (5.8) and (5.9), Theorem 5.1 shows that $f(x)$ has the representation (5.1) with $\varphi(t) \in L^p$; (5.2) is then an easy deduction.

6. Representation theorems: $\alpha(t)$ of bounded variation.

THEOREM 6.1. *A set of necessary and sufficient conditions for $f(x)$ to have the representation*

$$f(x) = \int_0^\infty e^{-xt} d\alpha(t), \qquad x > 0, \tag{6.1}$$

with $\alpha(t)$ of bounded variation on $(0, \infty)$, is

(i) *$f(x)$ continuous and bounded in $0 < x < \infty$;*

(ii) *for $k = 1, 2, \cdots$,*

$$\int_0^\infty |A_k(0, t)| dt \leqq M < \infty. \tag{6.2}$$

Condition (i) is trivially necessary; (ii) is necessary because by Lemma 3.1

$$\begin{aligned}\int_0^\infty |A_k(0, t)| dt &\leqq (2k - 1)!c_k \int_0^\infty t^{k-1} dt \int_0^\infty \frac{v^k}{(t + v)^{2k}} |d\alpha(v)| \\ &= (2k - 1)!c_k \int_0^\infty v^k |d\alpha(v)| \int_0^\infty \frac{t^{k-1}}{(t + v)^{2k}} dt \\ &= \theta_k \int_0^\infty |d\alpha(v)|,\end{aligned}$$

where $\theta_1 = 1$, $\theta_k = (k - 1)/k$ if $k > 1$.

To establish the sufficiency of the conditions, we appeal to Theorem 2.2. We obtain

$$f(x) - f(x + h) = \lim_{k \to \infty} \int_0^\infty e^{-xt}(1 - e^{-ht}) A_k(0, t)\, dt, \qquad 0 < x < \infty, \tag{6.3}$$

for each $h > 0$. The functions

[14] Loc. cit.

$$\alpha_k(t) = \int_0^t A_k(0, u)\,du$$

have, by (6.2), uniformly bounded variation on $(0, \infty)$. By Helly's theorem,[15] a subsequence of the $\alpha_k(t)$ converges to a function $\alpha(t)$ of bounded variation on $(0, \infty)$. Integrating by parts, we have

$$f(x) - f(x + h) = \lim_{k\to\infty} \int_0^\infty \alpha_k(t) \frac{\partial}{\partial t} \{e^{-xt}(1 - e^{-ht})\}\,dt.$$

By (6.2) and Lebesgue's limit theorem we may take the limit under the integral sign. A second integration by parts gives

$$f(x) - f(x + h) = \int_0^\infty e^{-xt}(1 - e^{-ht})\,d\alpha(t)$$

$$= \int_0^\infty e^{-xt}\,d\alpha(t) - \int_0^\infty e^{-(x+h)t}\,d\alpha(t) \qquad (0 < x < \infty;\, h > 0),$$

since each of the last integrals converges. As $h \to \infty$,

$$\int_0^\infty e^{-(x+h)t}\,d\alpha(t) \to \alpha(0+).$$

Hence $f(\infty)$ exists, and

$$f(x) = f(\infty) - \alpha(0+) + \int_0^\infty e^{-xt}\,d\alpha(t),$$

or

$$(6.4) \qquad f(x) = \int_0^\infty e^{-xt}\,d\alpha^*(t), \qquad x > 0,$$

where $\alpha^*(t)$ differs from $\alpha(t)$ only in having an additional jump of amount $f(\infty) - \alpha(0+)$ at $t = 0$.

We now derive from Theorem 6.1 the sufficiency part of the following known representation theorem.

THEOREM 6.2. *A necessary and sufficient condition for $f(x)$ to have the representation* (6.1) *with $\alpha(t)$ of bounded variation on $(0, \infty)$ is that $f(x)$ be of class C^∞ on $(0, \infty)$ and*

$$(6.5) \qquad \frac{1}{(k-1)!} \int_0^\infty u^{k-1} |f^{(k)}(u)|\,du \leqq M < \infty \qquad (k = 1, 2, \cdots).$$

From (6.5) with $k = 1$ we see that $f(0+)$ and $f(\infty)$ exist. It then follows from (6.5) that $f^{(k)}(x)x^k$ approaches a limit as x becomes infinite and as x approaches zero for $k = 0, 1, 2, \cdots$. Hence by the Hardy-Littlewood theorem cited in §5 we have

$$f^{(k)}(x) = o(x^{-k}) \qquad (x \to \infty,\, x \to 0+;\, k = 1, 2, \cdots).$$

[15] See, for example, Zygmund, *Trigonometrical Series*, Warszawa-Lwów, 1935, p. 80.

Consequently we may integrate by parts in the definition of $A_k(0, t)$, the integrated terms vanishing, so that

$$A_k(0, t) = -c_k(-t)^{k-1} \int_0^\infty u^{2k-1} e^{-tu} f^{(k)}(u)\, du,$$

$$\begin{aligned} \int_0^\infty | A_k(0, t) | \, dt &\leqq c_k \int_0^\infty t^{k-1}\, dt \int_0^\infty u^{2k-1} e^{-tu} | f^{(k)}(u) | \, du \\ &= c_k \int_0^\infty u^{2k-1} | f^{(k)}(u) | \, du \int_0^\infty t^{k-1} e^{-tu}\, dt \\ &= \frac{1}{k(k-2)!} \int_0^\infty u^{k-1} | f^{(k)}(u) | \, du \\ &\leqq M \qquad (k = 1, 2, \cdots). \end{aligned}$$

The result now follows from Theorem 6.1.

7. **Representation theorems: $\alpha(t)$ increasing.** We shall require an elementary Tauberian lemma.

LEMMA 7.1. *If $A(t) \geqq 0$ and*

$$\int_0^\infty (1 - e^{-ht}) A(t)\, dt \leqq M, \qquad h > 0,$$

then

$$\int_0^\infty A(t)\, dt \leqq M.$$

Since $1 - e^{-ht}$ is a positive increasing function of t, for $\epsilon > 0$ we have

$$\int_\epsilon^\infty A(t)\, dt \leqq \frac{1}{1 - e^{-h\epsilon}} \int_\epsilon^\infty (1 - e^{-ht}) A(t)\, dt \leqq \frac{M}{1 - e^{-h\epsilon}}.$$

Letting $h \to \infty$ and then $\epsilon \to 0$, we have

$$\int_\epsilon^\infty A(t)\, dt \leqq M, \qquad \int_0^\infty A(t)\, dt \leqq M.$$

THEOREM 7.2. *A necessary and sufficient condition for $f(x)$ to have the representation*

$$f(x) = \int_0^\infty e^{-xt}\, d\alpha(t), \qquad x > 0, \tag{7.1}$$

with $\alpha(t)$ non-decreasing and bounded on $0 < t < \infty$ is that $f(x)$ be non-negative, continuous, and bounded on $0 < x < \infty$, and

$$A_k(0, t) \geqq 0 \qquad (k = 1, 2, \cdots; 0 < t < \infty). \tag{7.2}$$

For the necessity, we have only to establish (7.2). We have, by Lemma 3.1,

$$A_k(0, t) = (2k - 1)!c_k t^{k-1} \int_0^\infty \frac{v^k}{(t + v)^{2k}} d\alpha(v) \geqq 0, \qquad k \geqq 1.$$

We now establish the sufficiency. By Theorem 2.2, we have for $x > 0$ and $h > 0$

$$f(x) - f(x + h) = \lim_{k \to \infty} I_k(x, h), \tag{7.3}$$

where

$$I_k(x, h) = \int_0^\infty e^{-xt}(1 - e^{-ht})A_k(0, t)\, dt$$

$$= (2k - 1)!c_k \int_0^\infty f(u)u^{k-1}\left\{\frac{x^k}{(x + u)^{2k}} - \frac{(x + h)^k}{(x + h + u)^{2k}}\right\} du.$$

Let $M > 0$ be an upper bound for $|f(x)|$ on $(0, \infty)$. If $k \geqq 2$ we have

$$0 \leqq I_k(x, h) \leqq (2k - 1)!c_k M \int_0^\infty u^{k-1}\left\{\frac{x^k}{(x + u)^{2k}} + \frac{(x + h)^k}{(x + h + u)^{2k}}\right\} du$$

$$= \frac{2(k - 1)M}{k} < 2M \qquad (h > 0).$$

Since $A_k(0, t) \geqq 0$, $I_k(x, h)$ is a decreasing function of x, $x > 0$. Since $I_k(x, h)$ is also bounded, $I_k(0+, h)$ exists. By a well-known trivial Tauberian theorem, we have

$$\int_0^\infty (1 - e^{-ht})A_k(0, t)\, dt = I_k(0+, h) < 2M.$$

By Lemma 7.1,

$$\int_0^\infty A_k(0, t)\, dt < 2M \qquad (k = 2, 3, \cdots). \tag{7.4}$$

We can now complete the proof either directly or by appealing to Theorem 6.1.

THEOREM 7.3. *In order that*

$$f(x) = \int_0^\infty e^{-xt}\, d\alpha(t), \qquad x > a, \tag{7.5}$$

with $\alpha(t)$ non-decreasing in $0 \leqq t < \infty$, it is necessary and sufficient that $f(x)$ be non-negative and continuous in $a < x < \infty$ and bounded in $(a + 1, \infty)$, and that

$$A_k(c, t) \geqq 0 \qquad (c > a; k = 1, 2, \cdots; 0 < t < \infty). \tag{7.6}$$

The necessity of the continuity and boundedness of $f(x)$ is easily established. To establish the necessity of (7.6), we consider the functions

$$f_\epsilon(x) = f(a + \epsilon + x), \qquad \epsilon > 0.$$

We may write

$$f_\epsilon(x) = \int_0^\infty e^{-xt}\, d\beta(t), \qquad x > -\epsilon,$$

with

$$\beta(t) = \int_0^t e^{-(\epsilon+a)u}\, d\alpha(u), \qquad t > 0,$$

$$\beta(0) = 0.$$

It is evident that $\beta(t)$ is non-decreasing; and $\beta(t)$ is bounded because $\beta(\infty) = f(a + \epsilon)$. Hence, by Theorem 7.2, the operator $A_k(0, t)$, formed for $f_\epsilon(x)$, is non-negative for $0 < t < \infty$, $k = 1, 2, \cdots$; that is,

$$c_k t^{k-1} \int_0^\infty Q_k(t, u) f(u + a + \epsilon)\, du \geqq 0.$$

But the left side, multiplied by $e^{(a+\epsilon)t}$, is $A_k(a + \epsilon, t)$, formed for $f(x)$. Thus (7.6) is established.

For the sufficiency, we consider again the functions $f_\epsilon(x)$, each of which is non-negative, continuous and bounded in $0 < t < \infty$. Since $A_k(0, t)$ for $f_\epsilon(x)$ becomes $A_k(\epsilon + a, t)$ for $f(x)$, if multiplied by $e^{(a+\epsilon)t}$, we can apply Theorem 7.2, and conclude that

$$f_\epsilon(x) = \int_0^\infty e^{-xt}\, d\alpha_\epsilon(t), \qquad x > 0,$$

where $\alpha_\epsilon(t)$ is non-decreasing and bounded in $(0, \infty)$. That is,

$$f(x) = \int_0^\infty e^{-xt} e^{(x+a)t}\, d\alpha_\epsilon(t)$$

$$= \int_0^\infty e^{-xt}\, d\beta_\epsilon(t), \qquad x > a + \epsilon,$$

where

$$\beta_\epsilon(0) = 0, \qquad \beta_\epsilon(t) = \int_0^t e^{(x+a)t}\, d\alpha_\epsilon(t),$$

so that $\beta_\epsilon(t)$ is non-decreasing in $0 < t < \infty$. Since $f(x)$ can have only one representation as a Laplace-Stieltjes integral, all the functions $\beta_\epsilon(x)$ must coincide (except possibly on a countable set of points, where they can be redefined without affecting the representation); the representation (7.5) follows at once.

We conclude this section by showing directly that the condition of Theorem 7.3 is equivalent to the condition that $f(x)$ is completely monotonic in $a < x < \infty$: that is, that

$$(7.7) \qquad (-1)^k f^{(k)}(x) \geqq 0 \qquad (a < x < \infty; k = 0, 1, 2, \cdots),$$

or that

$$\Delta_h^k f(x) \geqq 0 \qquad (a < x < \infty;\ h > 0;\ k = 0, 1, 2, \cdots). \tag{7.8}$$

Conditions (7.7) and (7.8) are known to be equivalent. In (7.8),

$$\Delta_h^0 f(x) = f(x), \qquad \Delta_h^k f(x) = \Delta_h^{k-1} f(x) - \Delta_h^{k-1} f(x + h) \qquad (k = 1, 2, \cdots).$$

THEOREM 7.4. *A function $f(x)$, defined in $a < x < \infty$, is completely monotonic there if and only if it satisfies the conditions imposed in Theorem* 7.3.

We show first that a function satisfying the hypotheses of Theorem 7.3 is completely monotonic. We again introduce the functions $f_\epsilon(x) = f(x + \epsilon + a)$, $\epsilon > 0$, and apply Theorem 2.3 to $f_\epsilon(x)$. We obtain

$$\Delta_h^n [f_\epsilon(x)] = \lim_{k\to\infty} \int_0^\infty \Delta_h^n [e^{-xt}] A_k(0, t)\, dt \tag{7.9}$$

for $n \geqq 1$, $\epsilon > 0$, $h > 0$, $0 < x < \infty$, where $A_k(0, t)$ is, of course, formed from $f_\epsilon(x)$. Since $A_k(0, t)$ for $f_\epsilon(x)$, when multiplied by $e^{(a+\epsilon)t}$, becomes $A_k(a + \epsilon, t)$ for $f(x)$, and since e^{-xt} is completely monotonic in $t > 0$ for each $x > 0$, it follows from (7.9) and the hypothesis (7.6) that $\Delta_h^n[f_\epsilon(x)] \geqq 0$ for $n \geqq 1$, $h > 0$, $\epsilon > 0$, $0 < x < \infty$. Since, by hypothesis, $f(x) \geqq 0$ (and hence $f_\epsilon(x) \geqq 0$), $f_\epsilon(x)$ satisfies (7.8). That is, $f_\epsilon(x)$ is completely monotonic in $0 < x < \infty$, $f(x)$ is completely monotonic in $a + \epsilon < x < \infty$ for every positive ϵ, and so $f(x)$ is completely monotonic in $a < x < \infty$.

We now suppose that we have a function $f(x)$ which is completely monotonic in $a < x < \infty$. Taking $c > a$, we have for $t > 0$ and $k = 1, 2, \cdots$,

$$A_k(c, t) = c_k t^{k-1} e^{ct} \int_0^\infty \frac{\partial^k}{\partial u^k} [u^{2k-1} e^{-tu}] f(u + c)\, du,$$

where the integral obviously exists because $f(u + c)$ is non-negative and decreasing. Since[16]

$$x^k f^{(k)}(x) \to 0 \qquad (x \to \infty;\ k = 1, 2, \cdots),$$

we may integrate by parts k times, obtaining

$$A_k(c, t) = c_k t^{k-1} e^{ct} \int_0^\infty u^{2k-1} e^{-tu} (-1)^k f^{(k)}(u + c)\, du;$$

by (7.7), $A_k(c, t) \geqq 0$ for $c > a$, $t > 0$, $k = 1, 2, \cdots$. This completes the proof.

8. **A Tauberian theorem.** We shall prove the following Tauberian theorem of Hardy and Littlewood.[17]

[16] R. P. Boas, *Asymptotic relations for derivatives*, this Journal, vol. 3(1937), pp. 637–646; Theorem $1B_2$, p. 638.

[17] G. H. Hardy and J. E. Littlewood, *Notes on the theory of series* (XI): *On Tauberian theorems*, Proceedings of the London Mathematical Society, (2), vol. 30(1929–30), pp. 23–37. We prove the result in the form involving a Stieltjes integral which seems to have been treated first by J. Karamata, *Neuer Beweis und Verallgemeinerung einiger Tauberian-Sätze*, Mathematische Zeitschrift, vol. 33(1931), pp. 294–299.

THEOREM 8.1. *If $\alpha(t)$ is non-decreasing,*

$$f(x) = \int_0^\infty e^{-xt}\, d\alpha(t), \qquad x > 0, \tag{8.1}$$

and

$$f(x) \sim Hx^{-r}, \qquad x \to 0+, \tag{8.2}$$

with $r > 0$ and $H > 0$, then

$$\alpha(t) \sim \frac{H}{\Gamma(r+1)} t^r, \qquad t \to \infty. \tag{8.3}$$

We break the proof into a number of lemmas.

LEMMA 8.2. *For $r > 0$ and $k = 1, 2, \cdots$,*

$$\Gamma(2k) \sum_{n=0}^{k} (-1)^n \binom{k}{n} \frac{\Gamma(2k-n-r-1)}{\Gamma(2k-n)} = (-1)^k \frac{\Gamma(k+r+1)\Gamma(k-r-1)}{\Gamma(r+1)}.$$

We have

$$\begin{aligned}\frac{\Gamma(k+r+1)\Gamma(k-r-1)}{\Gamma(2k)} &= \int_0^\infty \frac{u^{k+r}\,du}{(u+1)^{2k}} = \int_0^\infty \frac{u^r}{(u+1)^k}\left(1 - \frac{1}{u+1}\right)^k du \\ &= (-1)^k \sum_{n=0}^{k} (-1)^n \binom{k}{n} \int_0^\infty \frac{u^r\,du}{(u+1)^{2k-n}} \\ &= (-1)^k \sum_{n=0}^{k} (-1)^n \binom{k}{n} \frac{\Gamma(r+1)\Gamma(2k-n-r-1)}{\Gamma(2k-n)}.\end{aligned}$$

LEMMA 8.3. *If $g(t)$ is integrable in $(0, R)$ for every $R > 0$, if*

$$f(x) = \int_0^\infty e^{-xt} \varphi(t)\, dt,$$

the integral converging for some value of x, and if

$$g(t) \sim Ht^r, \qquad t \to 0+,$$

for some $r > 0$, $H \neq 0$, then

$$\int_0^\infty g(t) e^{-xt}\, dt \sim \frac{H\Gamma(r+1)}{x^{r+1}}, \qquad x \to \infty.$$

This is a well-known Abelian theorem.[18]

LEMMA 8.4. *If $\alpha(t)$ is positive and non-decreasing, $r > 0$, and for every sufficiently large k*

$$\int_0^\infty \frac{t^k \alpha(t)}{(x+t)^{2k}}\, dt \sim \frac{H}{\Gamma(r+1)} \frac{1}{x^{k-r-1}} \frac{\Gamma(k+r+1)\Gamma(k-r-1)}{(2k-1)!}, \qquad x \to \infty, \tag{8.4}$$

with $H > 0$, then

[18] Compare Hardy and Littlewood, p. 27 of their article cited in footnote 17.

$$\alpha(t) \sim \frac{H}{\Gamma(r+1)} t^r, \qquad t \to \infty. \tag{8.5}$$

This, except for changes in notation, was proved by Hardy and Littlewood. We reproduce the proof, arranged somewhat differently, for the sake of completeness.

In the first place,[19]

$$\alpha(t) = O(t^r), \qquad t \to \infty. \tag{8.6}$$

For, since $\alpha(t)$ increases and $t^k(x+t)^{-2k}$ increases if $t < x$, we have

$$\alpha(\tfrac{1}{2}x) \leqq \left(\frac{9x}{2}\right)^k \frac{2}{x} \int_{\frac{1}{2}x}^{x} \frac{t^k \alpha(t)\,dt}{(x+t)^{2k}} \leqq (3x)^k \frac{2}{x} \int_0^\infty \frac{t^k \alpha(t)\,dt}{(x+t)^{2k}} = O(x^r).$$

We now write (with $0 < \zeta < 1$)

$$I = \int_0^\infty \frac{t^k \alpha(t)\,dt}{(x+t)^{2k}} = \int_0^1 + \int_1^{(1-\zeta)x} + \int_{(1-\zeta)x}^{(1+\zeta)x} + \int_{(1+\zeta)x}^\infty$$
$$= I_0 + I_1 + I_2 + I_3,$$

and

$$J = \int_0^\infty \frac{t^{k+r}\,dt}{(x+t)^{2k}} = \int_0^{(1-\zeta)x} + \int_{(1-\zeta)x}^{(1+\zeta)x} + \int_{(1+\zeta)x}^\infty$$
$$= J_1 + J_2 + J_3.$$

We regard $\alpha(t)$ and r as fixed, and denote by C any constant; by $\epsilon_k(x)$ any function of k and x which approaches zero, for fixed k, as $x \to \infty$; by $\eta_\zeta(k)$ any function of ζ and k which approaches zero, for fixed ζ, as $k \to \infty$;[20] by $\lambda(\zeta)$ any function of ζ which approaches *unity* as $\zeta \to 0$.

We have

$$G(k) \equiv \frac{\Gamma(k+r+1)\Gamma(k-r-1)}{\Gamma(2k)} = k^{-\frac{1}{2}} 2^{-2k}(1+\eta(k)), \tag{8.7}$$

by Stirling's formula.

In our present notation, (8.4) states that

$$I = \frac{H + \epsilon_k(x)}{\Gamma(r+1)} G(k) x^{r-k+1}; \tag{8.8}$$

and (8.6), that

$$\alpha(t) < Ct^r, \qquad t \geqq 1. \tag{8.9}$$

Moreover, we have

$$J = G(k)x^{r-k+1} = k^{-\frac{1}{2}} 2^{-2k}(1+\eta(k))x^{r-k+1}. \tag{8.10}$$

[19] In the application we shall know (8.6) already.

[20] $\eta(k)$ denotes an $\eta_\zeta(k)$ which is constant with respect to ζ.

Since $t/(1+t)^2$ has a single maximum, at $t = 1$, where it has the value $\frac{1}{4}$, we have

$$x^{k-r-1}J_3 \leqq \left\{\frac{1+\zeta}{(2+\zeta)^2}\right\}^{k-r-2} \int_{1+\zeta}^{\infty} \frac{t^{2r+2}\,dt}{(1+t)^{2r+4}}$$
$$\leqq C2^{-2k}\{\sigma(\zeta)\}^{2k},$$

where $0 < \sigma(\zeta) < 1$. Thus, by (8.10),

$$x^{k-r-1}J_3 = \eta_\zeta(k)G(k). \tag{8.11}$$

Similarly,

$$x^{k-r+1}J_1 = \eta_\zeta(k)G(k). \tag{8.12}$$

Now

$$I_0 \leqq \alpha(1) \int_0^1 \frac{t^k\,dt}{(x+t)^{2k}} = \epsilon_k(x)x^{r-k+1}. \tag{8.13}$$

Using (8.9), we have

$$\begin{aligned} I_1 &\leqq C \int_0^{(1-\zeta)x} \frac{t^{k+r}\,dt}{(x+t)^{2k}} < CJ_1 \\ &= \eta_\zeta(k)G(k)x^{r-k+1}, \end{aligned} \tag{8.14}$$

by (8.12). Similarly,

$$I_3 = \eta_\zeta(k)G(k)x^{r-k+1}. \tag{8.15}$$

Since $\alpha(t)$ is non-decreasing,

$$\begin{aligned} I_2 = \int_{(1-\zeta)x}^{(1+\zeta)x} \frac{t^k\alpha(t)\,dt}{(x+t)^{2k}} &\leqq \frac{\alpha(x+\zeta x)}{(x-\zeta x)^r} \int_{(1-\zeta)x}^{(1+\zeta)x} \frac{t^{k+r}\,dt}{(x+t)^{2k}} \\ &= x^{-r}\alpha(x+\zeta x)\lambda(\zeta)J_2. \end{aligned} \tag{8.16}$$

Similarly,

$$I_2 \geqq x^{-r}\alpha(x-\zeta x)\lambda(\zeta)J_2. \tag{8.17}$$

From (8.16),

$$\begin{aligned} \frac{\alpha(x+\zeta x)}{(x+\zeta x)^r} &\geqq \frac{I_2}{J_2}\lambda(\zeta) \geqq \frac{I - I_0 - I_1 - I_3}{J}\lambda(\zeta) \\ &\geqq \left\{\frac{H+\epsilon_k(x)}{\Gamma(r+1)} + \epsilon_k(x) + \eta_\zeta(k)\right\}\lambda(\zeta), \end{aligned} \tag{8.18}$$

by successive use of (8.8), (8.13), (8.14), (8.15), (8.10).

Similarly, from (8.17),

$$\begin{aligned} \frac{\alpha(x-\zeta x)}{(x-\zeta x)^r} &\leqq \frac{I_2}{J_2}\lambda(\zeta) \leqq \frac{I}{J - J_1 - J_3}\lambda(\zeta) \\ &\leqq \frac{H+\epsilon(k)}{\Gamma(r+1)}\,\frac{\lambda(\zeta)}{1-\eta_\zeta(k)}, \end{aligned} \tag{8.19}$$

by (8.8), (8.10), (8.12), (8.11).

From (8.18) and (8.19), (8.5) follows: we choose ζ small, then k large, and finally let $x \to \infty$ for fixed ζ and k.

We now prove Theorem 8.1. Since

$$f(x) \sim Hx^{-r}, \qquad x \to 0+,$$

we have

$$f(x)x^{2k-n-2} \sim Hx^{2k-n-r-2}, \qquad x \to 0+.$$

For $k > r$ and $0 \leqq n \leqq k$, by Lemma 8.3,[21]

$$\int_0^\infty \frac{f(u)}{u} u^{2k-n-1} e^{-tu}\, du \sim H\Gamma(2k - n - r - 1)t^{n+r+1-2k}, \qquad t \to \infty.$$

Hence, since

$$\frac{\partial^k}{\partial u^k}[u^{2k-1}e^{-tu}] = e^{-tu}(-t)^k(2k-1)!\sum_{n=0}^{k}\binom{k}{n}\frac{u^{2k-n-1}}{(2k-n-1)!}(-t)^{-n},$$

$$\int_0^\infty \frac{f(u)}{u}\frac{\partial^k}{\partial u^k}[u^{2k-1}e^{-tu}]\,du$$

$$\sim (-1)^k Ht^{r-k+1}\Gamma(2k)\sum_{n=0}^{k}(-1)^n\binom{k}{n}\frac{\Gamma(2k-n-r-1)}{\Gamma(2k-n)} \tag{8.20}$$

$$= Ht^{r-k+1}\frac{\Gamma(k+r+1)\Gamma(k-r-1)}{\Gamma(r+1)}, \qquad t \to \infty,$$

by Lemma 8.2.

We may suppose without loss of generality that $\alpha(0) = 0$. We then have

$$\alpha(x) = \int_0^x d\alpha(t) \leqq e^{xy}\int_0^x e^{-yt}\,d\alpha(t) \leqq e^{xy}\int_0^\infty e^{-yt}\,d\alpha(t)$$

$$= e^{xy}f(y),$$

for any positive y. Taking $y = 1/x$, we get

$$\alpha(x) \leqq f(1/x) = O(x^r), \qquad x \to \infty. \tag{8.21}$$

Furthermore, we have

$$\frac{f(x)}{x} = \int_0^\infty e^{-xt}\alpha(t)\,dt,$$

and (8.21) shows that Lemma 3.1 is applicable to $f(x)/x$, so that for sufficiently large k

[21] The integral converges because $f(\infty) = \alpha(0+)$.

$$\int_0^\infty \frac{f(u)}{u} \frac{\partial^k}{\partial u^k} [u^{2k-1} e^{-tu}] \, du = (2k-1)! \int_0^\infty \frac{u^k \alpha(u) \, du}{(u+t)^{2k}}.$$

Consequently, from (8.20),

$$\int_0^\infty \frac{u^k \alpha(u) \, du}{(u+t)^{2k}} \sim \frac{Ht^{r-k+1}}{\Gamma(r+1)} \frac{\Gamma(k+r+1)\Gamma(k-r-1)}{\Gamma(2k)}, \qquad t \to \infty,$$

for large k. The desired result now follows from Lemma 8.4.

DUKE UNIVERSITY AND HARVARD UNIVERSITY.

THE ERGODIC THEOREM FOR A SEQUENCE OF FUNCTIONS

By Philip T. Maker

1. **Introduction.** The Birkhoff ergodic theorem,[1] for a single measure-preserving transformation $T(P)$ on a space Ω with a measure[2] m and a function $f(P)$ in $L_1(\Omega)$, states that

$$f^*(P) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f(T^n P)$$

exists a.e.[3] in Ω. In the case of a flow (i.e., a one-parameter family of measure-preserving transformations T_t for which $T_t \cdot T_s = T_{t+s}$ $(-\infty < t < \infty, -\infty < s < \infty)$, and for which, if A is measurable in Ω, the set of (P, t) points such that[4] $T_t P \epsilon A$ is measurable in the space $\Omega \times t$), the theorem states that $\lim_{T\to\infty} T^{-1} \int_0^T f(P_t)\, dt$ exists for almost all P in Ω.

The object of this paper is to extend the theorem to a dominated, convergent, double sequence of functions $\{f_{mn}(P)\}$ of $L_1(\Omega)$ for which $\lim_{m,n\to\infty} f_{mn}(P) = f_0(P)$, to obtain a more general ergodic theorem and a condition for $\lim_{m,n\to\infty} f^*_{mn}(P) = f^*_0(P)$. When $f_{mn}(P)$ is the characteristic function of a set E_{mn} and $T(P)$ is metrically transitive, the theorems allow the usual interpretation: that certain time means may be replaced by space means.

We assume throughout that Ω has a measure m defined on it and that $m(\Omega)$ is finite.

2. The discrete case.

Theorem 1. *If the complex-valued functions $f_{mn}(P)$ $(m, n = 1, 2, \cdots)$ and the positive function $\phi(P)$ are summable over Ω with $|f_{mn}(P)| < \phi(P)$, and $\lim_{m,n\to\infty} f_{mn}(P) = f_0(P)$ a.e. in Ω, then for any sequences of positive integers $\{m_{ij}\}$ and $\{n_{ij}\}$ for which $\lim_{i,j\to\infty} m_{ij} = \lim_{i,j\to\infty} n_{ij} = \infty$*

$$\lim_{i\to\infty} \frac{1}{i} \sum_{j=0}^{i-1} f_{m_{ij}n_{ij}}(T^j P) = f^*_0(P), \tag{1}$$

Received May 25, 1939; presented to the American Mathematical Society, April 8, 1939.

[1] G. D. Birkhoff, Proceedings of the National Academy of Science, vol. 17(1931), pp. 650–660. See also A. Khintchine, Mathematische Annalen, vol. 107(1933), pp. 485–488, and E. Hopf, *Ergodentheorie*, Berlin, 1937.

[2] A discussion of an abstract space with measure m is to be found in Hopf, loc. cit., p. 1.

[3] The abbreviation a.e. means almost everywhere.

[4] $T_t P$ will be written as P_t.

except for a set of measure zero independent of $\{m_{ij}\}$ *and* $\{n_{ij}\}$. *Moreover,*

$$(2) \qquad \lim_{m,n\to\infty} f^*_{mn}(P) = f^*_0(P) \quad a.e.$$

Proof. Suppose that (1) is false. It follows that, for each point P of a set S of measure $b > 0$,

$$\overline{\lim_{i\to\infty}} \frac{1}{i} \left| \sum_{j=0}^{i-1} f_{m_{ij}n_{ij}}(T^jP) - f^*_0(P) \right| > a > 0,$$

where the sequences $\{m_{ij}\}$ and $\{n_{ij}\}$ may depend on P. Hence from the definition of $f^*_0(P)$,

$$\overline{\lim_{i\to\infty}} \frac{1}{i} \left| \sum_{j=0}^{i-1} [f_{m_{ij}n_{ij}}(T^jP) - f_0(T^jP)] \right| > a.$$

We shall denote $| f_{m_{ij}n_{ij}}(T^jP) - f_0(T^jP) |$ by $g_{ij}(P)$, so that

$$\overline{\lim_{i\to\infty}} \frac{1}{i} \sum g_{ij}(P) > a \text{ on } S.$$

There exist positive numbers δ_r $(r = 1, 2, \cdots)$ such that, for every set E in Ω with $m(E) < \delta_r$, $\int_E \phi(P)\, dm < 2^{-r}$. By Egoroff's theorem, for each number δ_r there is an integer m_r so that for $m > m_r$, $n > m_r$, $| f_{mn}(P) - f_0(P) | < \frac{1}{2}a$ except in E_r, $m(E_r) < \delta_r$. For a fixed r, and a fixed P in S, denote those transforms T^jP which are in E_r by $T^{j'}P$, and the rest by $T^{j''}P$. Then

$$(3) \qquad a < \overline{\lim_{i\to\infty}} \frac{1}{i} \sum g_{ij}(P) \leqq \overline{\lim_{i\to\infty}} \frac{1}{i} \sum g_{ij'}(P) + \overline{\lim_{i\to\infty}} \frac{1}{i} \sum g_{ij''}(P).$$

But if i and j'' are greater than a sufficiently large integer n_r, then $m_{ij} > m_r$, $n_{ij} > m_r$, and $g_{ij''}(P) < \frac{1}{2}a$, so that

$$\frac{1}{i} \sum g_{ij''}(P) = \frac{1}{i} \sum_{j''\leqq n_r} g_{ij''}(P) + \frac{1}{i} \sum_{j''>n_r} g_{ij''}(P) < \frac{2B\cdot n_r}{i} + \tfrac{1}{2}a,$$

where

$$B = \max \{\phi(P), \phi(TP), \cdots, \phi(T^{n_r}P)\}.$$

Therefore $\overline{\lim}_{i\to\infty} i^{-1} \sum g_{ij''}(P) \leqq \frac{1}{2}a$, and so from (3)

$$\tfrac{1}{2}a < \overline{\lim_{i\to\infty}} \frac{1}{i} \sum g_{ij'}(P).$$

Define $\phi_r(P) = \phi(P)$ on E_r and 0 elsewhere for $r = 1, 2, \cdots$. Since

$$\overline{\lim_{i\to\infty}} \frac{1}{i} \sum_{j=0}^{i-1} 2\phi_r(T^jP) > \overline{\lim_{i\to\infty}} \frac{1}{i} \sum g_{ij'}(P) > \tfrac{1}{2}a,$$

we have

$$\phi_r^*(P) = \lim_{i\to\infty} \frac{1}{i} \sum_{j=0}^{i-1} \phi_r(T^jP) > \tfrac{1}{4}a,$$

for all P in S, except at most for a set of measure zero, and therefore[5]

$$\begin{aligned} \tfrac{1}{4}ab < \int_\Omega \phi_r^*(P)\, dm &= \lim_{i\to\infty} \int_\Omega \frac{1}{i} \sum_{j=0}^{i-1} \phi_r(T^jP)\, dm \\ &= \lim_{i\to\infty} \frac{1}{i} \sum_{j=0}^{i-1} \int_{T^j\Omega} \phi_r(P)\, dm = \int_{E_r} \phi(P)\, dm < \frac{1}{2^r}, \end{aligned}$$

and since this relation is impossible for every r, we have reached a contradiction and the proof of (1) is complete.

To prove (2), let P be any point for which (1) holds, and for which all means $f_{mn}^*(P)$ exist, and let $f_{m_kn_k}(P)$ $(k = 1, 2, \cdots)$ be any subsequence of $\{f_{mn}(P)\}$, convergent to $f_0(P)$ as $k \to \infty$. By the ergodic theorem, for any ϵ (>0), and all k, there are integers $i_k \to \infty$ such that

$$\left| \frac{1}{i} \sum_{j=0}^{i-1} f_{m_kn_k}(T^jP) - f_{m_kn_k}^*(P) \right| < \epsilon \qquad \text{for } i > i_k.$$

But by (1),

$$\left| \frac{1}{i} \sum_{j=0}^{i-1} f_{m_kn_k}(T^jP) - f_0^*(P) \right| < \epsilon$$

for all i greater than some sufficiently large i_0. Hence if k is so large that $i_k > i_0$, $|f_{m_kn_k}^*(P) - f_0^*(P)| < 2\epsilon$; since $\{f_{m_kn_k}(P)\}$ is any subsequence, the proof of (2) is complete.

3. The flow.

We now apply this theorem to the flow in the usual way.

THEOREM 2. *Let $f_\lambda(P)$ for $a \leqq \lambda \leqq b$ be a continuous family of functions of $L_1(\Omega)$ with $|f_\lambda(P)| < \phi(P)$, also in $L_1(\Omega)$. Let $\lambda(T, t)$ for $0 \leqq t \leqq T \leqq \infty$ be continuous with range in (a, b). Then $\lim_{T\to\infty} T^{-1} \int_0^T f_{\lambda(T,t)}(P_t)\, dt$ exists a.e. in Ω and equals $\lim_{T\to\infty} T^{-1} \int_0^T f_{\lambda_\infty}(P_t)\, dt$, where $\lambda_\infty = \lim_{T,t\to\infty} \lambda(T, t)$.*

Proof. According to the measurability condition in the definition of a flow, $f_\lambda(P_t)$, for each fixed value of λ, is a measurable function in the product-space

[5] The ergodic theorem also states that when $m(\Omega) < \infty$, $f^*(P)$ is in $L_1(\Omega)$ and

$$\lim_{N\to\infty} \int_\Omega N^{-1} \sum_{n=0}^{N-1} f(T^nP)\, dm = \int_\Omega f^*(P)\, dm.$$

$\Omega \times t$. Hence for each fixed pair, λ and P,[6] $f_\lambda(P_t)$ is measurable in t. Therefore, for a fixed T and P, $f_{\lambda(T,t)}(P_t)$ is measurable in t, because, as we immediately show, it is the limit of the measurable functions $\psi_n(t) = f_{\lambda(T,i/n)}(P_t)$ for $in^{-1}T \leqq t \leqq (i+1)n^{-1}T$ ($i = 0, 1, \cdots, n-1$, $n = 2, 3, \cdots$). For any particular t_0, if we denote by t_n the left-hand end-point of the interval in which t_0 is at the n-th subdivision, we have

$$\lim_{n\to\infty} \psi_n(t_0) = \lim_{n\to\infty} f_{\lambda(T,t_n)}(P_{t_0}) = f_{\lambda(T,t_0)}(P_{t_0}).$$

In addition to being measurable, $f_{\lambda(T,t)}(P_t)$ is dominated by $\phi(P_t)$, so that $\int_0^T f_{\lambda(T,t)}(P_t)\, dt$ exists for almost all P in Ω. Then, if n $(= [T])$ denotes the greatest integer less than T,

$$\begin{aligned} J &= \frac{1}{T}\int_0^T f_{\lambda(T,t)}(P_t)\, dt = \frac{1}{T}\int_0^n f_\lambda\, dt + \frac{1}{T}\int_n^T f_\lambda\, dt \\ &= J_1 + J_2. \end{aligned}$$

But $|J_2| < T^{-1}\int_n^T \phi(P_t)\, dt$. Since the limit of the latter as $T \to \infty$ is 0, $\lim_{T\to\infty} J_2 = 0$. Also

$$J_1 = T^{-1}\sum_{j=0}^{n-1}\int_0^1 f_{\lambda(T,t+j)}(P_{t+j})\, dt.$$

Let $\{T_i\}$ be any sequence of positive numbers such that $\lim_{i\to\infty} T_i = \infty$. The functions of P, $\int_0^1 f_{\lambda(T_i,t+j)}(P_t)\, dt$ ($i = 1, 2, \cdots$, $j = 0, 1, 2, \cdots, n_i - 1$) are dominated by $\int_0^1 \phi(P_t)\, dt$ and $\lim_{i,j\to\infty}\int_0^1 f_{\lambda(T_i,t+j)}(P_t)\, dt$ exists and is equal to $\int_0^1 f_{\lambda_\infty}(P_t)\, dt$, so that Theorem 1 applies to them. Hence

$$\lim_{i\to\infty} T_i^{-1}\sum_{j=0}^{n_i-1}\int_0^1 f_{\lambda(T_i,t+j)}(P_{t+j})\, dt$$

exists a.e. in Ω and equals $\lim_{T\to\infty} T^{-1}\int_0^T f_{\lambda_\infty}(P_t)\, dt$. Since $\{T_i\}$ is an arbitrary sequence with ∞ as its limit, we conclude that

$$\lim_{T\to\infty} J = \lim_{T\to\infty} J_1 = \lim_{T\to\infty}\frac{1}{T}\int_0^T f_{\lambda_\infty}(P_t)\, dt, \qquad \text{a.e. in } \Omega.$$

This completes the proof of Theorem 2.

RUTGERS UNIVERSITY.

[6] S. Saks, *Theory of the Integral*, New York, 1937, p. 83.

MONOTONE COVERINGS AND MONOTONE TRANSFORMATIONS

By A. D. Wallace

1. **Introduction.** The purpose of this paper is to relate coverings with connected sets and transformations whose inverse sets are connected. We shall use the term *covering* in a restricted sense. All coverings shall be with *closed sets* and the number of sets in a covering shall be *finite* unless it is explicitly stated that the sets arise as *inverse sets of points* in connection with a continuous transformation. It is also supposed that the term *space* shall mean *compact metric space* and the term *transformation* shall imply *continuity* and *single-valuedness*.

Let $C: A = \sum A_\alpha$ be a covering of the space A. We may associate with C an abstract complex called the *nerve*[1] of C in such a fashion that the vertices of this complex correspond in a one-to-one way with the sets of the covering. If we let $N(C) = K$ denote this correspondence, then the vertices $N(A_{\alpha_0}), \cdots, N(A_{\alpha_n})$ are spanned by an n-simplex if and only if the sets $A_{\alpha_0}, \cdots, A_{\alpha_n}$ have a non-vacuous product. We shall say that two coverings are *equivalent* if they are in one-to-one correspondence in such a way that the property of intersecting is preserved. The *dimension* of C is the largest integer n such that $n + 1$ different sets of C intersect. The words *chain, simple chain, simple closed chain,* and *acyclic* have their usual meanings and will be applied both to the covering and to its nerve. The covering C is said to be a *monotone covering* if the sets of C are *connected*. This terminology is by way of analogy with monotone transformation.

The next two definitions are generalizations of the notion of a fixed-point free transformation and are due to H. Hopf.[2] The covering $C: A = \sum A_\alpha$ is said to be *free* provided there exists a continuous transformation f of A into[3] itself such that $f(A_\alpha)A_\alpha = 0$ for each α. A transformation $T(A) = B$ is *free* if the covering of A with the sets $T^{-1}(y)$, $y \in B$, is free. This last is equivalent, as Hopf points out, to the condition that there exist a continuous transformation $f(A) \subset A$ such that $Tf(x) \neq T(x)$ for each $x \in A$. It is clear that if A does not admit a fixed-point free transformation, then it does not admit a free covering or a free transformation of any sort. Or, in other words, if every transformation of A into itself has a fixed-point, then A has no free covering and it is not

Received March 28, 1939; presented to the American Mathematical Society, December 27, 1938.

[1] Alexandroff-Hopf, *Topologie* I, Berlin, 1936, p. 152.

[2] H. Hopf, *Freie Überdeckungen und freie Abbildungen*, Fundamenta Mathematicae, vol. 28(1937), p. 31. All references to Hopf are to this paper. Both the methods and results of the present paper are intimately related to Hopf's work.

[3] If $T(A) \subset B$, then T maps A *into* B. We use *onto* if $T(A) = B$.

possible to define a free transformation on A. In case the space A admits a fixed-point free transformation into itself, it is of interest to determine approximately the greatest amount of freedom such a transformation is permitted. One way of doing this is to determine the type of free covering or free transformation that A admits or does not admit.

2. A characterization of multicoherence. It has been shown by Kuratowski[4] that in order that a locally connected continuum have the fixed-point property it is necessary that it be unicoherent.[5] Or, what is the same thing, if a locally connected continuum is multicoherent (that is, *not* unicoherent), there is a transformation of A into itself which leaves no point fixed. The following result is analogous:

THEOREM 1. *In order that a locally connected continuum be multicoherent it is necessary and sufficient that it admit a one-dimensional monotone covering that is not acyclic.*

Proof. That the condition is sufficient is Hopf's Satz I′. If A is not unicoherent, there exist a simple closed curve J in A and a transformation $r(A) = J$ which is such that $r(x) = x$ on J.[6] Let $J = t_1 + t_2 + t_3$ be a decomposition into three arcs having only their end-points in common. For each j let s_j denote an open arc of J containing t_j and such that $\bar{s}_1\bar{s}_2\bar{s}_3$ is vacuous. The components of $r^{-1}(s_j)$ are open and cover $A_j = r^{-1}(t_j)$. We select a finite collection covering A_j and let P_j denote the sum of these components. Let $A_{j1}, A_{j2}, \cdots, A_{jn_j}$ be the components of $\bar{P}_j$ and notice that if two sets $A_{j_1i_1}$ and $A_{j_2i_2}$ have anything in common then j_1 and j_2 are different or the two sets are identical. Since t_j is connected, it lies in some set A_{jk} and we adjust the notation so that it lies in A_{j1}. The covering $C: A = \sum A_{jk}$ is at least one-dimensional since A is connected and is monotone by construction. If three different sets of C intersect, then their first subscripts are different, as we noted above. If we assume that $A_{1k_1}A_{2k_2}A_{3k_3} \neq 0$, it easily follows that $\bar{s}_1\bar{s}_2\bar{s}_3 \neq 0$. For $P_j \subset r^{-1}(s_j)$ and hence $A_{jk_j} \subset \overline{r^{-1}(s_j)}$, so that, since A is compact,

$$r(A_{jk_j}) \subset r(\overline{r^{-1}(s_j)}) = \overline{rr^{-1}(s_j)} = \bar{s}_j.$$

Thus it follows that

$$0 \neq r(A_{1k_1}A_{2k_2}A_{3k_3}) \subset r(A_{1k_1})r(A_{2k_2})r(A_{3k_3}) \subset \bar{s}_1\bar{s}_2\bar{s}_3 .$$

Hence C is one-dimensional. Now since $t_it_j \neq 0$ we must have $A_{i1}A_{j1} \neq 0$, so that the sets A_{11}, A_{21}, A_{31} form a simple closed chain on the covering C.

The following somewhat intuitive result will be of use in a later theorem.

[4] C. Kuratowski, Fundamenta Mathematicae, vol. 14(1929), p. 304.

[5] A continuum is unicoherent if however it is expressed as the sum of two continua their product is again a continuum.

[6] K. Borsuk, Fundamenta Mathematicae, vol. 17(1931), p. 171.

LEMMA 1. *In order that a locally connected continuum be a dendrite it is necessary and sufficient that for each positive δ it admit a one-dimensional monotone acyclic δ-covering.*

Proof. Let A be a dendrite. It is well known[7] that for each positive δ there is a covering $C: A = \sum A_i$, where each A_i is a locally connected continuum of diameter less than δ, and such that no three sets intersect, the number of sets being finite. If the covering contained a simple closed chain, it would easily follow from the unique arc-wise connectivity that A contained a simple closed curve. The condition is thus necessary.

Assume that the locally connected continuum A contains a simple closed curve J, the sum of two arcs axb and ayb having only the points a and b in common. Let δ be a positive number less than one-third the distance from x to ayb and y to axb. Let C be a one-dimensional monotone δ-covering of A, and denote by C_j the subcollection of sets of C which intersect J. In C_j we can find chains C_{ax} and C_{xb} from a to x and x to b, respectively. Similarly we find chains C_{ay} and C_{yb}. We select from the first pair a chain C_x from a to b, one link of which contains x. There will be a similar chain C_y. Hence there will exist in C_j and therefore in C, two different simple chains from a to b. It easily follows that C contains a simple closed chain. Hence C is not acyclic.

3. **Monotone transformations.** The continuous transformation $T(A) = B$ is said to be *monotone*[8] provided that for each $y \in B$ the set $T^{-1}(y)$ is connected. If A is compact, the only case we consider, it is known that the inverse of every connected set is connected.

Hopf has shown (Satz C_1^*) that no continuum admits a free transformation into an arc. That is, if A is a continuum and $T(A) = B$, B a simple arc, then for any transformation $f(A) \subset A$ there will exist a point $x_0 \in A$ such that $Tf(x_0) = T(x_0)$. We extend this result by permitting the image to be a dendrite and at the same time restricting the transformation to be monotone. We discuss this in more detail after the following theorem.

THEOREM 2. *No continuum admits a free monotone transformation into a dendrite.*

Proof. Assume that the result is not true and let $T(A) = B$ be a free monotone transformation onto (since every subcontinuum of a dendrite is a dendrite) the dendrite B. By hypothesis there is a transformation $f(A) \subset A$ such that for each x we have $Tf(x) \neq T(x)$. Since A is compact, there is a positive δ such that for each $x \in A$, $\rho(Tf(x), T(x)) > \delta$. By Lemma 1, $B = \sum B_j$, $\delta(B_j) < \delta$, and this covering is one-dimensional, monotone and acyclic. The covering $C: A = \sum A_j$, $A_j = T^{-1}(B_j)$, has all of these properties since it is

[7] See, for example, K. Menger, *Kurventheorie*, Berlin, 1932, p. 191. We may also use Hopf's Hilfssatz 2 and Satz I′ to prove this.

[8] See G. T. Whyburn, American Journal of Mathematics, vol. 56(1934), p. 294. Also S. Eilenberg, Fundamenta Mathematicae, vol. 22(1934), p. 272.

clearly equivalent to the first one and since we know that the inverse of a connected set is connected. Now Hopf's Satz II′ is to the effect that no such covering can be free. Thus there is a j such that $f(A_j)A_j \neq 0$. But if $x \in A_j$, then $T(x) \in B_j$, and hence $Tf(x)$ is not in B_j (since $\delta(B_j) < \delta$ and $\rho(Tf(x), T(x)) > \delta$). It follows that $f(x)$ is not in A_j and thus that $f(A_j)A_j = 0$. This is a contradiction.

This theorem is not true if, for example, we replace "monotone" by the closely related "non-alternating"[9] as the following example shows: let J be the unit circle divided into three equal arcs by the points a, b, and c. Denote by z the origin (we may suppose that J is in the Euclidean plane) and let Y be the sum of the three straight line segments za, zb, and zc. Let $S(J) = Y$ be the transformation mapping the arc ab topologically onto the arc $az + zb$, bc onto $bz + zc$, and ca onto $cz + za$ so that the mid-points of the three arcs on J go into the origin. If then we put $U(J) = J$ for the transformation taking each point of J into its diametrically opposite, we see that (i) S is a non-alternating transformation, (ii) the set Y is a dendrite, (iii) for each $x \in J$ we have $SU(x) \neq S(x)$; that is, S is free. In essentially the same manner we can show that *if the locally connected continuum is multicoherent, it admits a free transformation onto a dendrite.* For if A is a multicoherent locally connected continuum, we may retract it onto the simple closed curve $J \subset A$, $r(A) = J$, $r(x) = x$, $x \in J$. Then if $f = Ur$, it follows that $rf(x) \neq r(x)$, $x \in A$, U as defined above, since there is no loss of generality in supposing that J is the unit circle. The transformation f then maps A onto $J \subset A$, and it may be seen that $SU(x) \neq S(x)$, $x \in J$, S as defined above. Then if we put $T = Sr$ and remember that $f = Ur$ and that r is the identity on J, we see that $Tf = SrUr = SUr \neq Sr = T$, and $T(A) = Y$, a dendrite.

From Theorem 2 we deduce the following well-known theorem of Scherrer:[10]

COROLLARY. *Every continuous transformation of a dendrite into itself admits a fixed point.*

Proof. Let $I(x) = x$ be the identity transformation on the dendrite A. Since I is monotonic, it is not free, so that if f maps A into itself, there is a point x_0 in A such that $If(x_0) = I(x_0)$, or $f(x_0) = x_0$.

LEMMA 2. *If $T(A) = B$ is free, and if T is factored in any way, $T = T_2T_1$, $T_1(A) = A'$, $T_2(A') = B$, then T_1 is also free.*

Proof. If T_1 is not free, there is a transformation $f(A) \subset A$ and a point x in A such that $T_1f(x) = T_1(x)$. Hence $T_2T_1f(x) = T_2T_1(x)$, and this is a contradiction.

[9] A transformation $T(A) = B$ is non-alternating if $A - T^{-1}(y) = A_1 + A_2$, A_1 and A_2 mutually separated, implies that $A_i = T^{-1}T(A_i)$. In connection with this paragraph see Hopf, pp. 49 and 54.

[10] W. Scherrer, Mathematische Zeitschrift, vol. 24(1926), p. 125.

THEOREM 3. *If a compact metric space admits a free transformation onto a set of dimension at most k, then it admits a free monotone transformation onto a set of dimension at most k.*

Proof. Let $T(A) = B$ be a free transformation, $\dim B \leqq k$. By a theorem of G. T. Whyburn (see footnote 8) we can factor T as indicated in Lemma 2 so that T_1 is monotone and $\dim A' \leqq \dim B \leqq k$. By the lemma cited T_1 is free.

From this result we can secure Hopf's Satz IIc:

COROLLARY. *No unicoherent locally connected continuum admits a free transformation onto a one-dimensional set.*

Proof. If it did, it would admit a free monotone transformation onto a one-dimensional set. But by Theorem 5 (which does not depend on this corollary) this implies that the image is a dendrite. Hence we would have a free monotone transformation into a dendrite in contradiction to Theorem 2.

For monotone transformations we have the following analogue of Hopf's Hilfssatz 3:

THEOREM 4. *A locally connected continuum admits a free monotone transformation onto an at most k-dimensional set if and only if it admits a free monotone covering of dimension at most k.*

Proof. Let $T(A) = B$ be a free monotone transformation, $\dim B \leqq k$, and let $f(A) \subset A$ be a transformation such that $Tf(x) \neq T(x)$. As in the proof of Theorem 2 there is a positive δ such that $\rho(Tf(x), T(x)) > \delta$. Since $\dim B \leqq k$, we can find a $\frac{1}{2}\delta$-covering of B with a finite number of closed sets and this covering is at most k-dimensional. By Hopf's Hilfssatz 2 we can replace this by a connected covering of the same dimension, $B = \sum B_j$, $\delta(B_j) < \delta$. Since T is monotone, we have an at most k-dimensional covering $A = \sum A_j$, $A_j = T^{-1}(B_j)$, with connected sets. The remainder of the argument proceeds as in the proof of Theorem 2. Thus the condition is necessary.

To show that the condition is sufficient we know that if A admits an at most k-dimensional covering with connected sets, then by Hopf's Hilfssatz 3 the space A admits a free transformation onto a set of dimension at most k. By Theorem 3 it admits a free monotone transformation onto a space of dimension at most k.

Either directly, with an almost identical proof, or using the above theorem, we can deduce the following:

COROLLARY. *A locally connected continuum admits an at most k-dimensional free covering if and only if it admits an at most k-dimensional free monotone covering.*

COROLLARY. *A locally connected continuum admits a free transformation into a space of dimension at most k if and only if it admits a free monotone transformation into a space of dimension at most k.*

This result can be deduced either directly or from Theorems 3 and 4. The next theorem shows the intimate relation existing between monotone coverings and monotone transformations in an even more striking fashion.

THEOREM 5. *In order that a locally connected continuum be unicoherent, each of the following conditions is both necessary and sufficient:*

(a) *every one-dimensional monotone image is acyclic;*

(b) *every one-dimensional monotone covering is acyclic.*

Proof. If $T(A) = B$ is monotone and $\dim B = 1$, then B is a dendrite. For if B contained a simple closed curve, we could, as in the proof of Lemma 1, find an arbitrarily small one-dimensional monotone covering that was not acyclic. The inverses of these sets form a similar covering of A and by Theorem 1 it would follow that A could not be unicoherent. If A is not unicoherent, it admits a retraction $r(A) = J$ onto a simple closed curve J. By Whyburn's factor theorem cited above we could factor this retracting transformation $r = r_2r_1$, so that $r_1(A) = A'$ is monotone and $\dim A' = 1$. Since r is topological on J, r_1 is also and thus A' contains a simple closed curve.

Part (b) is another form of Theorem 1. We can, however, prove (a) independently of Theorem 1 and get (b) from (a).

4. **An application.** We suppose that the space A is a locally connected continuum. Eilenberg has shown[11] that if A is unicoherent, so also is every interior[12] image of A. Eilenberg's proof makes use of properties of the space of transformations of A into the unit circle. Another proof can be gotten from a theorem of G. T. Whyburn[13] to the effect that the first Betti number of A cannot be increased under an interior transformation. We shall make use of Theorem 5(b).

THEOREM 6. *If A is a unicoherent locally connected continuum and $T(A) = B$ is interior, then B is unicoherent.*

Proof. We have only to show that no one-dimensional monotone covering of B can contain a simple closed chain. Assume the contrary and let $C': B = \sum B_i'$ be a one-dimensional monotone covering which contains a simple closed chain. Let s be less than one-half the Lebesgue number[14] of C' and denote by B_i the closure of the component of the s-neighborhood of B_i' which contains B_i'. Then B_i is a closed and connected set, and it is easily seen that the covering $C: B = \sum B_i$ is equivalent to C'. Moreover, each set in C has a non-vacuous interior. It follows by a theorem of G. T. Whyburn[15] that if A_{ij} is a component

[11] S. Eilenberg, Fundamenta Mathematicae, vol. 24(1935), p. 175.

[12] A transformation is interior if it maps open sets onto open sets. See S. Stoïlow, *Théorie des Fonctions Analytiques*, Paris, 1938.

[13] G. T. Whyburn, this Journal, vol. 4(1938), p. 1.

[14] Alexandroff-Hopf, loc. cit., p. 101.

[15] G. T. Whyburn, this Journal, vol. 3(1937), p. 370.

of $T^{-1}(B_i)$, then $T(A_{ij}) = B_i$. Let us suppose that the notation is so adjusted that

$$B_1, \quad B_2, \quad \cdots, \quad B_n, \quad B_1$$

is a simple closed chain in C. We consider the covering $C^1: A = \sum A_{ij}$, where each A_{ij} is a component of $T^{-1}(B_i)$. The number of sets in C^1 being finite by Whyburn's theorem cited above, it is easily seen that C^1 is a one-dimensional monotone covering of A. We shall show that C^1 contains a simple closed chain, contrary to Theorem 5(b). If A_{11} is any component of $T^{-1}(B_1)$, then since $B_1B_2 \neq 0$, some component, say A_{21}, of $T^{-1}(B_2)$ must intersect A_{11}. Continuing in this way, we get a chain

$$A_{11}, \quad A_{21}, \quad \cdots, \quad A_{n1}$$

such that each set intersects the following but there are no other intersections. The set A_{n1} need not intersect A_{11}; but it must intersect some component, say A_{12}, of $T^{-1}(B_1)$. Unless the chain has closed, we can find a component A_{22} of $T^{-1}(B_2)$ which intersects A_{12}. This process can continue so that we get a chain

$$A_{11}, \quad A_{21}, \quad \cdots, \quad A_{n1}, \quad A_{12}, \quad A_{22}, \quad \cdots, \quad A_{n2}.$$

The number of components in $T^{-1}(B_i)$ being finite, it is clear that we can get a simple closed chain in C^1.

THE UNIVERSITY OF VIRGINIA.

THEOREMS OF THE PICARD TYPE

By Olaf Helmer

1. **Introduction.** We shall denote the *order* and the *exponent of convergence* of an integral function $f(z)$ by

$$\rho = \operatorname{ord} f \qquad \text{and} \qquad \alpha = \exp f, \tag{1}$$

respectively. It is known that the latter does not exceed the former and is usually equal to it. In the case where

$$\exp f < \operatorname{ord} f \tag{2}$$

occurs, we shall say that the function $f(z)$ is *exceptional.*

Picard's original theorem that an integral function $f(z)$ does not omit more than one value has been replaced by the stronger theorem of Picard-Borel which, in the above terminology, can be formulated as follows: For an integral function $f(z)$ there is at most one constant c for which the function $f(z) - c$ is exceptional.

This theorem has, on the one hand, been generalized so as to apply to meromorphic functions.[1] On the other, it has been refined in various directions; in particular, the constant c has been replaced by a polynomial,[2] and even by an integral function whose order is smaller than that of $f(z)$.[3] It is the main object of this paper to prove another similar generalization of the Picard-Borel theorem that goes a little further. We shall consider pairs of integral functions $f(z)$, $g(z)$, and it will be our object to inquire into the number of integral functions $A(z)$ whose order is less than the larger of the orders of $f(z)$ and $g(z)$, for which the function $f(z) + A(z) \cdot g(z)$ is exceptional. In the special case where $g(z) = 1$ we obtain the previously considered cases cited above.

Of the theorems leading up to these results, the first is a precise statement concerning the order of the product of two integral functions, while the second and third deal with the order of an exponential form $A \cdot e^F + B \cdot e^G$, where A, B are integral functions and F, G are polynomials. These two latter theorems are not new, but have been proved in what follows for reasons of completeness.[4]

Throughout this paper we shall restrict ourselves to integral functions of *finite* order.

Received April 27, 1939.

[1] Cf. R. Nevanlinna, *Le Théorème de Picard-Borel*, Paris, 1929.

[2] Cf. E. Borel, *Leçons sur les Fonctions Entières*, 2d edition, Paris, 1921, p. 89.

[3] Cf. G. Valiron, *General Theory of Integral Functions*, Toulouse, 1923, p. 303.

[4] Theorem 3 has been used by Borel, but the proof he gives contains a mistake (see page 101 of reference in footnote 2). A correct proof can be found in G. Vivanti, *Theorie der eindeutigen analytischen Funktionen*, Leipzig, 1906.

2. Definitions. We begin by restating some of the fundamental definitions. Let $f(z)$ be an integral function with an m-fold zero at the origin ($m \geqq 0$) and its remaining zeros $z_1, z_2, z_3, \cdots$ numbered in such a way that $|z_1| \leqq |z_2| \leqq |z_3| \leqq \cdots$ (each zero, if necessary, being repeated according to its multiplicity). If $f(z)$ is of finite order, then it possesses a uniquely determined *Weierstrassian standard representation*:

$$f(z) = z^m \cdot e^{F(z)} \cdot \prod_i E\left(\frac{z}{z_i}, k\right), \tag{3}$$

where $F(z)$ is a polynomial (let its degree be q), k the smallest non-negative integer such that

$$\sum_i \frac{1}{|z_i|^{k+1}} \tag{4}$$

is convergent, and

$$E(x, k) = (1 - x) \cdot \exp\left(x + \frac{1}{2}x^2 + \cdots + \frac{1}{k}x^k\right). \tag{5}$$

The larger of the integers q and k is called the *genus* of $f(z)$ and is denoted by p:

$$p = \operatorname{gen} f = \max(q, k). \tag{6}$$

The smallest non-negative number α for which

$$\sum_i \frac{1}{|z_i|^{\alpha+\epsilon}} \qquad (\epsilon > 0) \tag{7}$$

converges, however small ϵ may be, is called the *exponent of convergence* of $f(z)$, written $\alpha = \exp f$.

If we put

$$M_f(r) = \max_{|z|=r} |f(z)|, \tag{8}$$

we may define the order ρ of $f(z)$, written $\rho = \operatorname{ord} f$, as the smallest non-negative number for which, however small ϵ,

$$M_f(r) < \exp r^{\rho+\epsilon} \qquad (\epsilon > 0), \tag{9}$$

provided r is large enough.

3. Lemmas presupposed in what follows. For later reference I shall now state the following lemmas; their proofs can be found in the works quoted above.

LEMMA 1. *If $f(z)$ and $g(z)$ are not identically zero, then*

$$\exp(f \cdot g) = \max(\exp f, \exp g). \tag{10}$$

LEMMA 2. *If $f(z)$ has the form* (3), $\rho = \operatorname{ord} f$, $\alpha = \exp f$, *and* $q = \deg F$, *then*

$$\rho = \max(q, \alpha). \tag{11}$$

The next three lemmas are immediate consequences of Lemma 2:

LEMMA 3.

$$\exp f \leqq \operatorname{ord} f. \tag{12}$$

LEMMA 4. *If* ord f *is non-integral, then*

$$\exp f = \operatorname{ord} f. \tag{13}$$

LEMMA 5. *Likewise,* (13) *holds for every canonical product, that is, for every function* $f(z)$ *of the form* (3), *where* $F(z)$ *is a constant.*

LEMMA 6.

$$\operatorname{ord}(f + g) \leqq \max(\operatorname{ord} f, \operatorname{ord} g). \tag{14}$$

Moreover,

$$\operatorname{ord} f \neq \operatorname{ord} g \rightarrow \operatorname{ord}(f + g) = \max(\operatorname{ord} f, \operatorname{ord} g). \tag{15}$$

LEMMA 7.

$$\operatorname{ord}(f \cdot g) \leqq \max(\operatorname{ord} f, \operatorname{ord} g). \tag{16}$$

Moreover, if $f(z)$ *and* $g(z)$ *are not identically zero, then*

$$\operatorname{ord} f \neq \operatorname{ord} g \rightarrow \operatorname{ord}(f \cdot g) = \max(\operatorname{ord} f, \operatorname{ord} g). \tag{17}$$

LEMMA 8. *Let* $f'(z)$ *denote the derivative of* $f(z)$*; then*

$$\operatorname{ord} f' = \operatorname{ord} f. \tag{18}$$

4. **The order of a product.** The information given by Lemma 7 will now be replaced by a more precise statement. If $f(z) = e^{F(z)}$ and $g(z) = e^{-F(z)}$, then, of course, the order of the product is less than the order of each factor. Apart from trivial exceptional cases of this type, however, the order of a product is always equal to the larger of the orders of the factors. Let, in general, $f(z)$ be of the form (3) (now writing k_f instead of k), and similarly

$$g(z) = z^n \cdot e^{G(z)} \cdot \prod_j E\left(\frac{z}{z_j}, k_g\right). \tag{19}$$

Then we can make the following assertion:

THEOREM 1.

$$\operatorname{ord}(f \cdot g) = \max(\operatorname{ord} f, \operatorname{ord} g) \tag{20}$$

unless all the following conditions are fulfilled:

$$\begin{cases} \text{(i)}\ \deg F > \exp f, & \text{(iii)}\ \deg F = \deg G, \\ \text{(ii)}\ \deg G > \exp g, & \text{(iv)}\ \deg(F + G) < \deg F, \end{cases} \tag{21}$$

in which case

$$\text{ord}\,(f\cdot g) < \text{ord}\,f = \text{ord}\,g. \tag{22}$$

Proof.

$$f\cdot g = z^{m+n}\cdot e^{F+G}\cdot\prod_i E\left(\frac{z}{z_i}, k_f\right)\cdot\prod_j E\left(\frac{z}{z_j'}, k_g\right). \tag{23}$$

If the conditions (21) are fulfilled, then by Lemma 2

$$\text{ord}\,f = \text{ord}\,g = \deg F = \deg G. \tag{24}$$

Hence, applying Lemma 7 to (23), we have

$$\begin{aligned}\text{ord}\,(f\cdot g) &\leqq \max\,(\text{ord}\,z^{m+n}, \text{ord}\,e^{F+G}, \text{ord}\,\textstyle\prod_i, \text{ord}\,\prod_j)\\ &< \max\,(1, \deg F, \deg F, \deg G) = \text{ord}\,f = \text{ord}\,g,\end{aligned} \tag{25}$$

and this proves (22). On the other hand, if

$$\text{ord}\,(f\cdot g) < \max\,(\text{ord}\,f, \text{ord}\,g), \tag{26}$$

then by Lemma 7, (17), we must have $\text{ord}\,f = \text{ord}\,g$, and by Lemmas 1 and 3 $\exp f < \text{ord}\,f$ and $\exp g < \text{ord}\,g$; but these, by virtue of Lemma 2, are precisely the conditions (i), (ii) and (iii) under (21). But then condition (iv) must also hold, since otherwise (23) would, by Lemma 7, have the same order as f, contrary to (26).

5. **The order of exponential forms.** Of the following two theorems the first is a special case of the second.

THEOREM 2. *Let $F(z)$ and $G(z)$ be polynomials; then*

$$\text{ord}\,(e^F - e^G) = \max\,(\deg F, \deg G) \tag{27}$$

unless $e^F \equiv e^G$, in which case, of course, $\text{ord}\,(e^F - e^G) = 0$.

Proof. Let

$$e^{F(z)} - e^{G(z)} = h(z), \tag{28}$$

and let, say, $\deg F \geqq \deg G$; assuming $\text{ord}\,h < \deg F$, we have to prove e^F and e^G to be identical. Differentiating (28), we get

$$F'e^F - G'e^G = h', \tag{29}$$

and eliminating e^G from (28) and (29), we obtain

$$(G' - F')\cdot e^F = G'h - h'. \tag{30}$$

The right side of (30), by Lemmas 6, 7, and 8, has an order less than $\deg F$, whereas the left side would have order $\deg F$ if $G' - F'$ were not identically zero; hence $G = F + c$, so that

$$e^F(1 - e^c) = h. \tag{31}$$

Again comparing orders, we find that $1 - e^c = 0$, so that $h \equiv 0$.

THEOREM 3. *Let $F(z)$ and $G(z)$ be polynomials, and $A(z)$ and $B(z)$ integral functions with*

$$\max(\operatorname{ord} A, \operatorname{ord} B) < \max(\deg F, \deg G). \tag{32}$$

Then

$$\operatorname{ord}(Ae^F + Be^G) = \max(\deg F, \deg G) \tag{33}$$

unless $Ae^F + Be^G \equiv 0$, *in which case* $\operatorname{ord}(Ae^F + Be^G) = 0$.

Proof. If F and G are not of the same degree, the assertion follows immediately from Lemmas 6 and 7. Hence we may assume $\deg F = \deg G = q$, say. Also let $\operatorname{ord} A \leqq \operatorname{ord} B$.

If A and B have any zeros in common, let

$$A = A^* \cdot D \quad \text{and} \quad B = B^* \cdot D, \tag{34}$$

where A^* and B^* are relatively prime, and where D can be chosen so that $\operatorname{ord} D \leqq \operatorname{ord} A$. Then

$$Ae^F + Be^G = D \cdot h, \tag{35}$$

where

$$A^*e^F + B^*e^G = h. \tag{36}$$

If we now suppose that $\operatorname{ord}(Ae^F + Be^G) < q$, we must also have $\operatorname{ord} h < q$ since $\operatorname{ord} D \leqq \operatorname{ord} A < q$. Differentiating (36), we obtain

$$(A^{*\prime} + A^*F')e^F + (B^{*\prime} + B^*G')e^G = h'. \tag{37}$$

Elimination of e^F from (36) and (37) leads to

$$e^G[(A^{*\prime} + A^*F')B^* - (B^{*\prime} + B^*G')A^*] = h(A^{*\prime} + A^*F') - h'A^*. \tag{38}$$

On comparing orders, we see that the factor of e^G must vanish identically, so that

$$A^*(B^{*\prime} + B^*G' - B^*F') = A^{*\prime}B^*. \tag{39}$$

A^* must, therefore, be a divisor of $A^{*\prime}B^*$, but A^* and B^* being relatively prime, this implies that A^* is a divisor of $A^{*\prime}$; this is only possible if A^* has no zeros, for if A^* and $A^{*\prime}$ had a zero in common, its multiplicity with respect to A^* would have to be larger by one than its multiplicity with respect to $A^{*\prime}$. The same argument holds good for B^*, so that we may put

$$A^* = e^U \quad \text{and} \quad B^* = e^V, \tag{40}$$

where U and V are polynomials of degree less than q. We now can apply Theorem 2 to

$$e^{U+F} + e^{V+G} = h \tag{41}$$

and find that h must be identically zero.

6. Theorems of the Picard type. A function $f(z)$ was called exceptional if its order exceeds its exponent of convergence. It must then, by Lemma 2, be of the form

$$f(z) = e^{F(z)} \cdot f_1(z), \tag{42}$$

where $\operatorname{ord} f_1 < \operatorname{ord} f$; it follows, in particular, that $\operatorname{ord} f = \deg F$, so that an exceptional function of finite order must of necessity be of an integral order.

We shall now consider pairs of integral functions $f(z)$, $g(z)$ and examine the functions of the form $f(z) + A(z)g(z)$ as to exceptionality, where the order of A is less than the larger of the orders of f and g. It will turn out that, as a rule, there is at most one function A for which $f + Ag$ is exceptional. Only when f and g are of the same order, there may be more than one such function A, and even then, apart from a trivial exception, the number of these functions A is restricted to at most two. The case where f and g have different orders is simpler and will be treated first.

THEOREM 4. *Let $f(z)$ and $g(z)$ be integral functions of distinct finite orders; then there is at most one integral function $A(z)$ with* $\operatorname{ord} A < \max(\operatorname{ord} f, \operatorname{ord} g)$ *for which the function $f(z) + A(z)g(z)$ is exceptional.*

Proof. Suppose both $f + Ag$ and $f + Bg$ are exceptional:

$$f + Ag = Ue^X \quad \text{and} \quad f + Bg = Ve^Y, \tag{43}$$

where

$$\max(\operatorname{ord} U, \operatorname{ord} V) < \max(\operatorname{ord} f, \operatorname{ord} g) = \deg X = \deg Y \tag{44}$$

(so that, in particular, $\max(\operatorname{ord} f, \operatorname{ord} g)$ must be an integer). Eliminating either f or g from (43), we have

$$(A - B)g = Ue^X - Ve^Y \quad \text{and} \quad (A - B)f = AVe^Y - BUe^X. \tag{45}$$

Using the first or the second of these equations, according as $\operatorname{ord} f > \operatorname{ord} g$ or $\operatorname{ord} f < \operatorname{ord} g$, we find by applying Theorem 3

$$A = B. \tag{46}$$

As a by-product we obtain the following

COROLLARY. *If $f(z)$ and $g(z)$ are integral functions of distinct finite orders, and if $\rho = \max(\operatorname{ord} f, \operatorname{ord} g)$ is non-integral, then there is no integral function $A(z)$ of an order less than ρ for which $f(z) + A(z)g(z)$ is exceptional.*

We now come to the case where $f(z)$ and $g(z)$ are of the same order. In this case we cannot expect quite so simple a result as Theorem 4, as the following examples show. (i) Let $f(z) = g(z) = e^{H(z)}$; then $f + Ag = (1 + A)e^H$ will be exceptional for every $A \neq -1$ with $\operatorname{ord} A < \deg H$. (ii) Let $f(z) = \cos z$ and $g(z) = \sin z$; then in $A = i$ and $A = -i$ we have two functions for which $f + Ag$ is exceptional.

The second example shows that even if we do not have the trivial exceptional situation of example (i), there may be two functions A with the property under consideration. However, the following theorem shows that, in general, the situation is at least no worse than has to be expected on account of the above examples.

THEOREM 5. *Let $f(z)$ and $g(z)$ be integral functions of the same finite order ρ. Then*

(a) *if f and g are multiples of the same exponential function:*

$$f(z) = e^{H(z)}f_1(z) \quad \text{and} \quad g(z) = e^{H(z)}g_1(z), \tag{47}$$

where $\operatorname{ord} f_1 < \rho$ *and* $\operatorname{ord} g_1 < \rho$, *every function* $f(z) + A(z)g(z)$ *with* $\operatorname{ord} A < \rho$ *is exceptional or identically zero;*

(b) *in every other case there are at most two integral functions $A(z)$ with* ord $A < \rho$ *for which the function $f(z) + A(z)g(z)$ is exceptional.*

Proof. The proof will proceed along the following lines. If we assume that there are three functions A with the property under consideration, it will follow that both f and g must themselves be exceptional. For a pair of exceptional functions f, g it will then be shown that there cannot be more than two functions A unless we have the situation described under (a).

Suppose, then, that the three functions $f + Ag$, $f + Bg$ and $f + Cg$ are exceptional:

$$f + Ag = Ue^X, \qquad f + Bg = Ve^Y, \qquad f + Cg = We^Z, \tag{48}$$

where A, B, C, U, V, W are of order less than ρ and A, B, C are distinct. Eliminating f and g from (48), we obtain

$$(A - C)Ve^{Y-X} + (A - B)We^{Z-X} = (B - C)U. \tag{49}$$

If $U \equiv 0$, we have $f = -Ag$ and hence $(B - A)g = Ve^Y$, so that both f and g are exceptional. If $U \not\equiv 0$, we may apply Theorem 3 to (49), with the result that $Y - X$ and $Z - X$ (and hence also $Y - Z$) must be of degree less than ρ. Let, say,

$$X = Y + X_1 = Z + X_2, \tag{50}$$

where $\deg X_1 < \rho$ and $\deg X_2 < \rho$. From (48) and (50) we have

$$VW(f + Ag) = UVWe^X = UWe^{X_1}(f + Bg) = UVe^{X_2}(f + Cg), \tag{51}$$

and hence

$$f\cdot(We^{X_1} - Ve^{X_2}) = g\cdot(CVe^{X_2} - BWe^{X_1}), \tag{52}$$

where both expressions in parentheses are of order less than ρ. Now let $d(z)$ be the greatest common divisor of $f(z)$ and $g(z)$ (d is unique apart from an exponential factor). On account of (48), d must be a divisor of U, so that

$$\exp d \leqq \exp U \leqq \operatorname{ord} U < \rho. \tag{53}$$

d can therefore be chosen so that

$$\text{ord } d < \rho. \tag{54}$$

Let

$$f = d \cdot f^* \quad \text{and} \quad g = d \cdot g^*, \tag{55}$$

where f^* and g^* are relatively prime and, on account of (54), of order ρ. Then (52) becomes

$$f^* \cdot S = g^* \cdot T. \tag{56}$$

Here S and T are not zero (because otherwise we would have $B = C$) and are of order less than ρ. It follows that f^* is a divisor of T, and likewise g^* a divisor of T, and therefore

$$\exp f^* \leqq \exp T < \rho \quad \text{and} \quad \exp g^* \leqq \exp S < \rho. \tag{57}$$

From (54) and (55), on the other hand, we see that $\exp f = \exp f^*$ and $\exp g = \exp g^*$; both, by (57), are less than ρ, so that f and g must be exceptional.

If f and g are exceptional, they must be of the form

$$f = e^F \cdot f_1 \quad \text{and} \quad g = e^G \cdot g_1 \tag{58}$$

with $\text{ord } f_1 < \rho$ and $\text{ord } g_1 < \rho$. Substituting in (52), we obtain

$$e^{F-G} f_1 \cdot (We^{X_1} - Ve^{X_2}) = g_1 \cdot (CVe^{X_2} - BWe^{X_1}). \tag{59}$$

The expression on the left is not identically zero; hence, comparing orders, we find $\deg (F - G) < \rho$, so that f and g are multiples of the same exponential function e^G:

$$f = e^G \cdot f_2 \ (\text{with } f_2 = e^{F-G} f_1) \quad \text{and} \quad g = e^G \cdot g_1 \tag{60}$$

with $\text{ord } f_2 < \rho$ and $\text{ord } g_1 < \rho$, and this is exactly the situation described under (a). In this case it is obvious that for every function A of order less than ρ the function $f + Ag$ is exceptional or zero.

Again we can state a corollary referring to the case where ρ is non-integral.

COROLLARY. *If $f(z)$ and $g(z)$ are integral functions of the same finite non-integral order ρ, then there is at most one integral function $A(z)$ of order less than ρ for which $f(z) + A(z)g(z)$ is exceptional.*

The possible existence of one such function A is shown by the following example. Let f be any function of a non-integral order greater than 1, and let $g = e^z - f$; then for $A = 1$ we have $f + Ag = e^z$ which is exceptional.

Proof of the Corollary. Supposing that the first two equations of (48) hold, we obtain

$$(A - B)g = Ue^X - Ve^Y. \tag{61}$$

Here, if $A - B$ were not zero, the order on the left would be ρ which is supposedly not an integer, whereas the order on the right, by Theorem 3, is max (deg X, deg Y) which is an integer. Hence B cannot be distinct from A.

It should be noticed that, for the function $f + Ag$ to be exceptional, it is necessary that

$$\exp (f + Ag) < \rho, \tag{62}$$

where $\rho = \max (\operatorname{ord} f, \operatorname{ord} g)$. In fact, in the case of Theorem 4 where ord $f \neq$ ord g and hence $\operatorname{ord} (f + Ag) = \rho$, this condition is also sufficient. But even when proving Theorem 5, where $\operatorname{ord} f = \operatorname{ord} g = \rho$, we have used nothing but (62) in assuming that $f + Ag$ be exceptional. This justifies the following additional

COROLLARY. *Theorems 4 and 5 remain valid when the words "is exceptional" are replaced by "has an exponent of convergence less than ρ", where* $\rho = \max$ $(\operatorname{ord} f, \operatorname{ord} g)$.

In view of Theorems 4 and 5 the obvious question arises as to the possible exceptionality of functions of the form $Af + Bg$. The answer is given in the following two theorems.

THEOREM 6. *Let $f(z)$ and $g(z)$ be integral functions of distinct finite orders. Then there is essentially at most one pair of integral functions $A(z)$, $B(z)$, both not identically zero, with* $\max (\operatorname{ord} A, \operatorname{ord} B) < \max (\operatorname{ord} f, \operatorname{ord} g)$, *for which the function $A(z)f(z) + B(z)g(z)$ is exceptional; essentially in the sense that every other pair $A_1(z)$, $B_1(z)$ must be proportional to the first.*

Proof. We may assume that ord $f >$ ord g. Suppose there are two pairs A, B and A_1, B_1 for which $Af + Bg$ and $A_1f + B_1g$ are exceptional:

$$Af + Bg = Ue^X \quad \text{and} \quad A_1f + B_1g = Ve^Y, \tag{63}$$

where A, B, A_1, B_1, U, V are of order less than that of f. Now let

$$f^* = AA_1f. \tag{64}$$

Then we have ord $f^* =$ ord f. From (63) and (64) we obtain

$$f^* + A_1Bg = A_1Ue^X \quad \text{and} \quad f^* + AB_1g = AVe^Y. \tag{65}$$

These equations, by Theorem 4, are incompatible unless $A_1B = AB_1$; in other words, the pair A_1,B_1 has to be proportional to the pair A,B.

THEOREM 7. *Let $f(z)$ and $g(z)$ be integral functions of the same finite order ρ. Then, in the case* (a) *of Theorem 5, every function $A(z)f(z) + B(z)g(z)$ with* $\max (\operatorname{ord} A, \operatorname{ord} B) < \max (\operatorname{ord} f, \operatorname{ord} g)$ *is exceptional or identically zero. In every other case, there are essentially at most two pairs of integral functions $A(z)$, $B(z)$, both not identically zero, with* $\max (\operatorname{ord} A, \operatorname{ord} B) < \max (\operatorname{ord} f, \operatorname{ord} g)$, *for which the function $A(z)f(z) + B(z)g(z)$ is exceptional, essentially in the sense that every other pair must be proportional to one of the first two pairs.*

Proof. In the trivial case (a) equations (47) hold, and we have

$$Af + Bg = e^{H}(Af_1 + Bg_1), \tag{66}$$

which is exceptional or identically zero. Otherwise suppose there are three pairs A,B, A_1,B_1, and A_2,B_2 for which $Af + Bg$, $A_1f + B_1g$ and $A_2f + B_2g$ are exceptional:

$$Af + Bg = Ue^{X}, \qquad A_1f + B_1g = Ve^{Y}, \qquad A_2f + B_2g = We^{Z}, \tag{67}$$

where A, B, A_1, B_1, A_2, B_2, U, V, W are of order less than ρ. Now let

$$f^* = AA_1A_2f. \tag{68}$$

Then we have $\operatorname{ord} f^* = \operatorname{ord} f$, and from (67) and (68) we obtain

$$\begin{gathered} f^* + A_1A_2Bg = A_1A_2Ue^{X}, \qquad f^* + AA_2B_1g = AA_2Ve^{Y}, \\ f^* + AA_1B_2g = AA_1We^{Z}. \end{gathered} \tag{69}$$

Applying Theorem 5 (b), we find that of the functions A_1A_2B, AA_2B_1 and AA_1B_2 at most two can be distinct. If, say, the last two coincide:

$$AA_2B_1 = AA_1B_2, \tag{70}$$

we find, since $A \not\equiv 0$, that the pair A_2,B_2 must be proportional to the pair A_1,B_1.

THE UNIVERSITY OF ILLINOIS.

A CLASS OF CONTINUED FRACTIONS

By H. M. Schwartz

Introduction. The algebraic continued fractions

$$\frac{k_1 z\,|}{|\,1} + \frac{k_2 z\,|}{|\,1} + \cdots, \qquad k_i \text{ complex and } \neq 0, \tag{1}$$

whose coefficients satisfy the condition

$$\sum_{i=1}^{\infty} |k_i| < \infty \tag{2}$$

possess the following interesting property:

The numerators and denominators of the approximants of (1) *form respectively two sequences that converge uniformly over any bounded region of the z-plane* [3].[1]

A more general class than (1) is given by

$$\frac{k_1\ \,|}{|\,z - c_1} + \frac{k_2\ \,|}{|\,z - c_2} + \cdots, \qquad k_i \text{ and } c_i \text{ complex}, \qquad k_i \neq 0. \tag{3}$$

In fact, except for simple changes of the variable and the fraction, (1) can be obtained from (3) by taking all $c_i = 0$ ([4], §61). In this paper we study some consequences of condition (2) for this general class of continued fractions.

The general case of unrestricted c_i is considered briefly in §1. A generalization of the above-mentioned property of (1) under (2) in this case is given in

Theorem 1. *Denote the n-th approximant of* (3) *by* $P_n(z)/Q_n(z)$. *If we have* (2), *then the two sequences*

$$\frac{P_n(z)}{\prod_{i=1}^{n}(z - c_i)}, \qquad \frac{Q_n(z)}{\prod_{i=1}^{n}(z - c_i)} \qquad (n = 1, 2, \cdots) \tag{4}$$

converge each uniformly in every domain[2] *of the z-plane at a positive distance from* $\{c_i\}$*—the set of points* c_i $(i = 1, 2, \cdots)$.

More precise results are obtained in §§2 and 3 for a number of special cases. Convergence properties of (3) are discussed in §4. In §5 we consider the special case corresponding to the condition

$$\sum_{i=1}^{\infty} \frac{1}{|c_i|} < \infty. \tag{5}$$

Received June 3, 1939; in revised form, November 10, 1939; presented to the American Mathematical Society, April 8, 1939, under a different title. The author wishes to express his gratitude to Professor J. A. Shohat for his encouragement and advice during the preparation of this paper, and for his valuable suggestions.

[1] Numbers in brackets refer to the bibliography.

[2] By *domain* we mean a point set in the extended complex plane.

In this case (2) can be replaced by the weaker condition

$$\sum_{i=1}^{\infty} \left| \frac{k_{i+1}}{c_i c_{i+1}} \right| < \infty. \tag{6}$$

Corresponding to Theorem 1 we have

THEOREM 2. *If the coefficients of* (3) *satisfy conditions* (5) *and* (6), *then the sequences*

$$\frac{P_n(z)}{\prod_{i=1}^{n} (-c_i)}, \quad \frac{Q_n(z)}{\prod_{i=1}^{n} (-c_i)} \qquad (n = 1, 2, \cdots) \tag{7}$$

converge each uniformly over any bounded domain of the z-plane, the limit functions being, thus, entire functions.

The second part of the paper applies to Bessel's continued fraction results obtained in the first part. The related system of Lommel polynomials is shown to form an orthogonal set [5] for certain values of the parameter and the corresponding weight function is found.

I. A class of continued fractions

1. **Proof of Theorem 1.** Since[3]

$$P_n(z) = k_1 Q_{n-1,1}(z) \qquad (n = 1, 2, \cdots), \tag{8}$$

it suffices to consider only $\{Q_n(z)\}$.

Let D be a given domain of the z-plane at the distance d (>0) from the set $\{c_i\}$. By the Euler-Minding formulas

$$\begin{aligned} Q_n(z) = \prod_{i=1}^{n} (z - c_i) \cdot \Bigg[1 &+ \sum_{i}^{1,n-1} \frac{k_{i+1}}{(z - c_i)(z - c_{i+1})} \\ &+ \sum_{i<j}^{1,n-2} \frac{k_{i+1}}{(z - c_i)(z - c_{i+1})} \frac{k_{j+2}}{(z - c_{j+1})(z - c_{j+2})} + \cdots \Bigg] \end{aligned} \tag{9}$$

we have for all z in D

$$\begin{aligned} \left| \frac{Q_n(z)}{\prod_{i=1}^{n} (z - c_i)} \right| &\equiv | V_n(z) | \leqq 1 + \sum_{i}^{1,n-1} \frac{|k_{i+1}|}{d^2} + \sum_{i<j}^{1,n-2} \frac{|k_{i+1}|}{d^2} \frac{|k_{j+2}|}{d^2} + \cdots \\ &\leqq \prod_{i=2}^{n} \left(1 + \frac{|k_i|}{d^2}\right) \leqq \prod_{i=2}^{\infty} \left(1 + \frac{|k_i|}{d^2}\right) = K, \text{ say,} \end{aligned}$$

the infinite product being convergent by assumption (2). Now by the recurrence relations

$$Q_n(z) = (z - c_n) Q_{n-1}(z) + k_n Q_{n-2}(z) \qquad (n = 2, 3, \cdots) \tag{10}$$

[3] $Q_{n,m}(z)$ $(m = 0, 1, \cdots)$ denotes, in Perron's notation ([4], §5) the function arising from $Q_n(z)$ when k_i and c_i are replaced by k_{i+m} and c_{i+m} respectively.

we find

$$V_n(z) = V_{n-1}(z) + \frac{k_n}{(z - c_n)(z - c_{n-1})} V_{n-2}(z) \quad (n = 2, 3, \cdots).$$

It follows that for all z in D

$$| V_n(z) - V_{n-1}(z) | \leqq \frac{|k_n|}{d^2} K,$$

and so the series

$$V_1(z) + [V_2(z) - V_1(z)] + \cdots$$

converges uniformly in D; that is, the sequence $\{V_n(z)\}$ converges uniformly in D.

A necessary condition. When the set $\{c_i\}$ is unbounded, it is easy to see by consulting Theorem 2 that condition (2) in Theorem 1 is not necessary for the validity of Theorem 1. Indeed, by (5) it follows, as a consequence of Theorem 2, that the sequences

$$\frac{P_n(z)}{\prod_{i=1}^{n} (z - c_i)} = \frac{P_n(z)}{\prod_{i=1}^{n} (-c_i)} \frac{1}{\prod_{i=1}^{n} \left(1 - \frac{z}{c_i}\right)}, \qquad \frac{Q_n(z)}{\prod_{i=1}^{n} (z - c_i)} = \frac{Q_n(z)}{\prod_{i=1}^{n} (-c_i)} \frac{1}{\prod_{i=1}^{n} \left(1 - \frac{z}{c_i}\right)}$$

converge uniformly in every bounded domain of the z-plane at a positive distance from $\{c_i\}$, while (6) shows that the set $\{k_i\}$ could even be unbounded.

Consider, then, the case of bounded $\{c_i\}$. By (9) we have for all z at a sufficiently great distance from the origin, the following expansion in ascending powers of z^{-1}:

$$V_n(z) = 1 + \sum_{i}^{1,n-1} k_i \left(\frac{1}{z} + \frac{c_i}{z^2} + \cdots\right)\left(\frac{1}{z} + \frac{c_{i+1}}{z^2} + \cdots\right) = 1 + \frac{1}{z^2} \sum_{i=2}^{n} k_i + \cdots.$$

By Weierstrass' double series theorem, it follows, therefore, that *when* $\{c_i\}$ *is bounded, then for the conclusion of Theorem* 1 *to hold true, it is necessary that* $\sum k_i$ *converge.*

Note. In the important special case

(11) $\qquad k_1 > 0, \quad k_i < 0 \quad (i = 2, 3, \cdots); \quad c_i \text{ real} \quad (i = 1, 2, \cdots)$

which is related to the moments problem of Stieltjes and Hamburger ([4], §72) it is thus seen that condition (2) in Theorem 1 in case of bounded $\{c_i\}$ is both necessary and sufficient.

2. When in addition to (2) it is also assumed that

(12) $$\lim_{i\to\infty} c_i \text{ exists and } = c, \text{ say},$$

it is possible to obtain a more precise result than that given in Theorem 1. We make use of the following

LEMMA. *If the sequence of analytic functions $f_n(z)$ $(n = 1, 2, \cdots)$ satisfies the following two conditions:*

(a) *$f_n(z)$ $(n = 1, 2, \cdots)$ is everywhere regular, except possibly at a given point c;*

(b) *there exists a sequence of closed contours C_n $(n = 1, 2, \cdots)$ enclosing the point c, with*[4]

$$\lim_{n\to\infty} \max_{z \text{ on } C_n} |z - c| = 0$$

and such that on each C_n, the sequence $\{f_n(z)\}$ converges uniformly;
then

(c) *sequence $\{f_n(z)\}$ converges uniformly in every domain at a positive distance from the point c, the limit function being an entire function in $(z - c)^{-1}$.*

The lemma is an immediate consequence of a well-known theorem of Weierstrass, which states that a sequence of functions regular analytic inside and on a given closed contour, and converging uniformly on the contour, converges uniformly in any closed domain in the interior of the domain bounded by the contour. We have only to subject the z-plane to the linear transformation

$$z' = \frac{1}{z - c}$$

and to observe that the functions

$$\phi_n(z') = f_n\left(\frac{1}{z'} + c\right) = f_n(z)$$

are, by condition (a) of the lemma, everywhere regular in the finite z'-plane, and that, according to condition (b) of the lemma, any finite domain of the z'-plane can be enclosed in a closed contour (namely, the map in the z'-plane of one of the contours C_n given in (b)) on which the sequence converges uniformly. By the above theorem, it follows that $\{\phi_n(z')\}$ converges uniformly in any bounded domain and its limit function, therefore, is an entire function. Translating this result back into the z-plane, we obtain the conclusion (c) of the lemma.

Consider now the sequences (4) of Theorem 1. It will be sufficient to consider one of them, say the second. By condition (12) and Theorem 1 (condition (2) being assumed), it can be readily concluded that we can find a sequence of contours C_n (each C_n being at some positive distance from the set $\{c_i\}$) which satisfy the conditions given in (b), with

$$f_n(z) = \frac{Q_n(z)}{\prod_{i=1}^{n} (z - c_i)}.$$

[4] In our application, it is sufficient to take for C_n a circle with c for center and radius which $\to 0$ as $n \to \infty$.

Suppose now that the set of integers $\{p_i\}$ $(i = 1, 2, \cdots)$ is such that

$$(13) \qquad \sum_{i=1}^{\infty} \left| \frac{c_i - c}{z - c} \right|^{p_i+1} \text{ converges for every } z \neq c$$

(it being always possible to determine such sets of integers, since by (12) $\lim_{i\to\infty} (c_i - c) = 0$), so that the infinite product

$$\prod_{i=1}^{\infty} \left(1 - \frac{c_i - c}{z - c}\right) \exp\left[\frac{c_i - c}{z - c} + \frac{1}{2}\left(\frac{c_i - c}{z - c}\right)^2 + \cdots + \frac{1}{p_i}\left(\frac{c_i - c}{z - c}\right)^{p_i}\right]$$

converges uniformly on each of the above contours C_n. By the identity

$$(14) \qquad \frac{Q_n(z)}{\prod_{i=1}^{n} (z - c_i)} = \frac{e^{T(z;n)}}{(z - c)^n} \frac{Q_n(z)}{e^{T(z;n)} \prod_{i=1}^{n} \left(1 - \frac{c_i - c}{z - c}\right)},$$

where

$$(15) \qquad T(z; n) = \sum_{i=1}^{n} \left[\frac{c_i - c}{z - c} + \frac{1}{2}\left(\frac{c_i - c}{z - c}\right)^2 + \cdots + \frac{1}{p_i}\left(\frac{c_i - c}{z - c}\right)^{p_i}\right],$$

it follows that condition (b) of the lemma holds also for the sequence

$$(16) \qquad \frac{e^{T(z;n)}}{(z - c)^n} Q_n(z) \qquad (n = 1, 2, \cdots).$$

But this sequence satisfies also assumption (a) of the lemma, the functions $Q_n(z)$ being polynomials in z. Hence, for sequence (16) conclusion (c) of the lemma is valid and we obtain

THEOREM 3. *If the coefficients of the continued fraction* (3) *satisfy conditions* (2) *and* (12), *then sequence* (16) *and sequence*

$$\frac{e^{T(z;n)}}{(z - c)^n} P_n(z) \qquad (n = 1, 2, \cdots),$$

where $T(z; n)$ are given by (15), *converge each uniformly in every domain of the z-plane at a positive distance from the point c, the limit functions being entire functions in $(z - c)^{-1}$.*

COROLLARY. *If we can take $p_i = p$ $(i = 1, 2, \cdots)$ in* (13), *and if we have further*

$$(17) \qquad \sum_{i=1}^{\infty} (c_i - c)^k \textit{ converges for } k = 1, 2, \cdots, p,$$

then the two sequences

$$(18) \qquad \left\{\frac{P_n(z)}{(z - c)^n}\right\}, \qquad \left\{\frac{Q_n(z)}{(z - c)^n}\right\}$$

converge uniformly in every domain at a positive distance from the point c.

In fact, we have only to note that by virtue of assumption (17), the following limit exists:

$$\lim_{n\to\infty} T(z; n) = \sum_{i=1}^{\infty} \sum_{s=1}^{p} \frac{1}{s} \left(\frac{c_i - c}{z - c}\right)^s,$$

this limit being, clearly, uniform with respect to z in every domain at a positive distance from the point c.

If instead of (12) we suppose that

$$\lim_{i\to\infty} \frac{1}{|c_i|} = 0, \tag{19}$$

then, by examining the proof of Theorem 3, it is easy to see that the argument of that proof can be retained. In this case, we have to use Weierstrass' theorem itself rather than the lemma. Moreover, in place of identity (14) we have to consider the following identity

$$\frac{Q_n(z)}{\prod_{i=1}^{n} (z - c_i)} = \frac{e^{R(z;n)} Q_n(z)}{\prod_{i=1}^{n} (-c_i)} \frac{1}{e^{R(z;n)} \prod_{i=1}^{n} \left(1 - \frac{z}{c_i}\right)},$$

where

$$R(z; n) = \sum_{i=1}^{n} \left[\frac{z}{c_i} + \frac{1}{2}\left(\frac{z}{c_i}\right)^2 + \cdots + \frac{1}{p_i}\left(\frac{z}{c_i}\right)^{p_i}\right], \tag{20}$$

the integers p_i being determined so that

$$\sum_{i=1}^{\infty} \left|\frac{z}{c}\right|^{p_i+1} \text{ converges for every finite } z. \tag{21}$$

THEOREM 4. *If the coefficients of the continued fraction* (3) *satisfy conditions* (2) *and* (19), *then each of the two sequences*

$$\left\{\frac{e^{R(z;n)}}{\prod_{i=1}^{n} (-c_i)} P_n(z)\right\}, \qquad \left\{\frac{e^{R(z;n)}}{\prod_{i=1}^{n} (-c_i)} Q_n(z)\right\},$$

where $R(z; n)$ are given by (20) *and* (21), *converges uniformly in every bounded domain of the z-plane, the limit functions being thus entire functions.*

If we have further $p_i = p$ $(i = 1, 2, \cdots)$ and

$$\sum_{i=1}^{\infty} \frac{1}{c_i^k} \textit{ converges for } k = 1, 2, \cdots, p,$$

then the above convergence behavior belongs already to the sequences (7).

3. The condition

$$\sum_{i=1}^{\infty} |c_i - c| < \infty \tag{22}$$

is included as the special case $p = 0$ in the corollary to Theorem 3. This case, however, can be proved directly by very simple means.

First, we restate it, for the sake of convenience, as

THEOREM 5. *If the coefficients of* (3) *satisfy conditions* (2) *and* (22), *then each of the sequences* (18) *converges uniformly in every domain of the z-plane at a positive distance from the point c.*

Writing for the sake of brevity $\frac{Q_n(z)}{(z-c)^n} = Q_n^*(z)$ $(n = 1, 2, \cdots)$, we find by the recurrence formulas (10)

$$Q_n^*(z) = \left(1 + \frac{c - c_n}{z - c}\right) Q_{n-1}^*(z) + \frac{k_n}{(z-c)^2} Q_{n-2}^*(z) \qquad (n = 2, 3, \cdots).$$

Therefore, for all z in a given domain D at the positive distance d from the point c,

$$| Q_n^*(z) - Q_{n-1}^*(z) | \leqq A \cdot \left(\frac{|c - c_n|}{d} + \frac{|k_n|}{d^2}\right) \quad (n = 2, 3, \cdots),$$

where

$$A = \underset{\substack{z \text{ in } D \\ n=1,2,\cdots}}{\bar{B}} \{| Q_n^*(z) |\}.$$

Moreover, by formulas (9), we have in D

$$\begin{aligned} Q_n^*(z) = \prod_{s=1}^{n} \left(1 - \frac{c_s - c}{z - c}\right) &+ \frac{1}{(z-c)^2} \sum_{i}^{1,n-1} k_{i+1} \prod_{s \neq i, i+1}^{1,n} \left(1 - \frac{c_s - c}{z - c}\right) \\ &+ \frac{1}{(z-c)^4} \sum_{i<j}^{1,n-2} k_{i+1} k_{j+2} \prod_{s \neq i, i+1, j+1, j+2}^{1,n} \left(1 - \frac{c_s - c}{z - c}\right) + \cdots. \end{aligned} \tag{23}$$

Consequently, by assumptions (2) and (22), we have for all z in D and for all n

$$\begin{aligned} | Q_n^*(z) | &\leqq \prod_{i=1}^{n} \left(1 + \frac{|c_i - c|}{d}\right) \left[1 + \sum_{i}^{1,n-1} \frac{|k_{i+1}|}{d^2} + \sum_{i<j}^{1,n-2} \frac{|k_{i+1}|}{d^2} \frac{|k_{j+2}|}{d^2} + \cdots\right] \\ &\leqq \prod_{i=1}^{n} \left(1 + \frac{|c_i - c|}{d}\right) \prod_{i=1}^{n} \left(1 + \frac{|k_i|}{d^2}\right) \\ &\leqq \prod_{i=1}^{\infty} \left(1 + \frac{|c_i - c|}{d}\right) \prod_{i=1}^{\infty} \left(1 + \frac{|k_i|}{d^2}\right). \end{aligned} \tag{24}$$

It follows that A is a finite number, and the proof of Theorem 5 may be completed exactly as the proof of Theorem 1.

The corollary to Theorem 3 shows that assumption (22) is not necessary for the validity of Theorem 5. However, the weaker assumption

$$\sum_{i=1}^{\infty} (c_i - c) \text{ converges} \tag{25}$$

is indeed *necessary*. In fact, we have

THEOREM 6. *If one of the sequences* (18), *say* $Q_n^*(z)$, *converges uniformly in some region D of the z-plane at a positive distance from the point c, then we must have* (25).

This follows by the application of Weierstrass' double series theorem to the expansion (cf. (23))

$$Q_n^*(z) = 1 + \frac{1}{z-c}\sum_{i=1}^{n}(c-c_i) + \frac{1}{(z-c)^2}\left[\sum_{i<j}^{1,n}(c-c_i)(c-c_j) + \sum_{i=2}^{n}k_i\right] + \cdots.$$

We note further that the existence of the limit for $n \to \infty$ of the coefficient

$$\sum_{i<j}^{1,n}(c-c_i)(c-c_j) + \sum_{i=2}^{n}k_i \tag{26}$$

of $(z-c)^{-2}$ in the above expansion will imply the condition

$$\sum_{i=1}^{\infty}k_i \text{ converges} \tag{27}$$

if we assume (22), since

$$\sum_{i<j}^{1,n}|(c-c_i)(c-c_j)| \leqq \left(\sum_{i=1}^{\infty}|c-c_i|\right)^2.$$

It follows, therefore, that

when the coefficients of fraction (3) *are such that the convergence of the series* $\sum k_i$, $\sum (c-c_i)$ *implies their absolute convergence (as might happen, for instance, when we have the conditions* (11)), *then the conditions* (2) *and* (22) *of Theorem* 5 *are necessary as well as sufficient for the validity of that theorem.*

Since

$$\sum_{i=1}^{n}(c-c_i)^2 = \left[\sum_{i=1}^{n}(c-c_i)\right]^2 - 2\sum_{i<j}^{1,n}(c-c_i)(c-c_j),$$

it is seen by Theorem 6 and by the form of the coefficient (26) that

if we assume (2) *or only* (27), *then the assumption in Theorem* 6 *implies in addition to* (25) *also*

$$\sum_{i=1}^{\infty}(c-c_i)^2 \textit{ converges.}$$

On the order of the limit functions of Theorem 5. It will be again sufficient to confine our attention to one of the sequences (18), say to $\{Q_n^*(z)\}$. Let

$$\lim_{n\to\infty} Q_n^*(z) = Q(z). \tag{28}$$

By (24), we find, r being an arbitrary positive number,

$$\max_{|z-c|=r}|Q(z)| \leqq \prod_{i=1}^{\infty}\left(1 + \frac{|c_i - c|}{r}\right)\prod_{i=1}^{\infty}\left(1 + \frac{|k_i|}{r^2}\right).$$

Hence, if we denote by ρ_1 and ρ_2 the exponents of convergence of the two series

$$\sum (c_i - c), \qquad \sum k_i$$

respectively (the exponent of convergence of a series $\sum_{i=1}^{\infty} a_i$ being defined as the greatest lower bound, if it exists, of the set of numbers $\{k\}$ for which $\sum_{i=1}^{\infty} | a_i |^k$ converges), it follows by a familiar reasoning in the theory of entire functions (namely, that concerning the order of a canonical product) that given an arbitrarily assigned positive number ϵ, we can find a positive number r_ϵ such that for $r \leqq r_\epsilon$ we have

$$\max_{|z-c|=r} | Q(z) | \leqq \exp \left[\left| \frac{1}{z - c} \right|^{\rho^* + \epsilon} \right],$$

where

$$\rho^* = \max \{\rho_1 , 2\rho_2\}.$$

By the definition of the order of an entire function, we have thus

THEOREM 7. *The order ρ of either the limit function* (28) *or the limit function*

$$P(z) = \lim_{n\to\infty} \frac{P_n(z)}{(z - c)^n}$$

of Theorem 5 *satisfies the inequality*

$$\rho \leqq \max \{\rho_1 , 2\rho_2\},$$

where ρ_1 and ρ_2 are defined above. In the special case

$$c_i = c \qquad (i = 1, 2, \cdots)$$

therefore, we have[5]

$$\rho \leqq 2\rho_2 . \tag{29}$$

In view of (2) *and* (22), *we have in all cases*

$$\rho \leqq 2.$$

In the particular case (11), which, as was mentioned before, is related to the moments problem of Stieltjes-Hamburger, it is easy to obtain additional information concerning the limit functions $P(z)$ and $Q(z)$, by taking in account the known fact that the zeros of $P_n(z)$ and $Q_n(z)$ are in that case all real ([4], §69; [5]), and by referring to the following theorem of Laguerre:[6]

The limit of a sequence of polynomials having only real zeros, which converges uniformly in every bounded domain of the z-plane, is an entire function of genus 1 *at most, except for a possible factor of the form e^{az^2}, where a is real.*

[5] A similar result for fractions (1) under (2) is given in [3].

[6] Oeuvres, vol. 1, pp. 174–177.

Since $P_n(z)(z-c)^{-n}$ and $Q_n(z)(z-c)^{-n}$ are (in case (11)) polynomials in $(z-c)^{-1}$ having only real zeros, we obtain immediately

THEOREM 8. *When the coefficients of the continued fraction* (3), *in addition to satisfying the assumptions given in Theorem* 5, *also satisfy* (11), *then the limit functions* $P(z)$ *and* $Q(z)$ *of Theorem* 5 *are each at most of genus* 1 *in* $(z-c)^{-1}$ *except for a possible factor of the form*

$$\exp\left[a\left(\frac{1}{z-c}\right)^2\right],\ a \text{ real}.$$

The case $c_i = c$. The zeros of $Q_n(z)$ are, in this case, symmetric about the point c,[7] and it is seen that

$$Q_{2n}(z) = \prod_{i=1}^{n}[(z-c)^2 - z_{i,2n}^2], \qquad Q_{2n+1}(z) = (z-c)\prod_{i=1}^{n}[(z-c)^2 - z_{i,2n+1}^2]$$

($c \pm z_{i,n}$ are the zeros of $Q_n(z)$),

so that

$$\frac{Q_n(z)}{(z-c)^n} = \prod_{i=1}^{[\frac{1}{2}n]}\left(1 - \frac{z_{i,n}^2}{(z-c)^2}\right).$$

It follows that the limit function $Q(z)$ is even, considered as a function of $(z-c)$. In the special case (11), it is then seen by Theorem 8 that $Q(z)$ is of genus 0 in $(z-c)^{-2}$ except for a possible factor of the form $\exp[a(z-c)^{-2}]$, a real. By utilizing also the separability properties of the zeros of $Q_n(z)$ under (11) [5], it can be shown independently of Laguerre's theorem that in this case $Q(z)$ *is exactly of genus* 0 *in* $(z-c)^{-2}$ (*i.e.*, $a = 0$).[8]

If we do not assume (11), then the above result can still be retained provided we suppose that ρ_2 of Theorem 7 is $\leqq \frac{1}{2}$. Indeed, this latter assumption implies by (29) that the order of the limit function $Q(z)$ is $\leqq 1$. Hence, since the limit function is even in $(z-c)$, we have

THEOREM 9. *The limit function* $Q(z)$ *of Theorem* 7 *will be of genus* 0 *in* $(z-c)^{-2}$ *when* $c_i = c$ *and when* ρ_2 *given in Theorem* 7 *is* $\leqq \frac{1}{2}$.

4. By Theorem 3, it is immediately apparent that the continued fraction (3), with the coefficients satisfying (2) and (12), converges at every point of the z-plane with the possible exception of the point c and of the zeros of the limit function of sequence (16); and that it converges uniformly in every domain of the z-plane at a positive arbitrarily small distance from the point c and from those zeros. By Hurwitz's theorem concerning the zeros of the limiting function of a convergent sequence of analytic functions [2] it is readily seen that the

[7] By means of the recurrence formulas (10) (and $Q_0 = 1$, $Q_1 = z - c_1$) we find

$$Q_n(c+z) = (-1)^n Q_n(c-z) \qquad (n = 0, 1, 2, \cdots).$$

[8] Cf. [6], p. 529, where a similar result is obtained for (1) under (2).

zeros in question are given by the limit points of the set of zeros of the polynomials $Q_n(z)$ $(n = 1, 2, \cdots)$, points which are zeros for an infinity of $Q_n(z)$ being considered as limit points.[9]

Convergence of (3) *under the assumption of* §3. We make use of the following two theorems:

THEOREM 10. *The limit functions $P(z)$ and $Q(z)$ of Theorem 5 have no common zeros.*[10]

THEOREM 11. *Given an arbitrarily assigned positive ϵ, we can find a number n_ϵ such that for all $n \geqq n_\epsilon$ the zeros of the polynomials $Q_n(z)$ of Theorem 5, which lie outside the circle $|z - c| = \epsilon$ do not coincide with any zeros of the limit function $Q(z)$.*

Theorem 10 can be proved in a manner analogous to Perron's proof of Maillet's theorem ([4], p. 346, footnote). In fact, writing (the existence of the limits is obviously assured by Theorem 5)

$$\lim_{n\to\infty} \frac{Q_{n,m}(z)}{(z-c)^n} = q_m(z) \qquad (m = 0, 1, \cdots; z \neq c),$$

we can show that

$$q_m(z) = \frac{z - c_{m+1}}{z - c} q_{m+1}(z) + \frac{k_{m+2}}{(z-c)^2} q_{m+2}(z) \quad (m = 0, 1, \cdots), \tag{30}$$

and

$$\lim_{m\to\infty} q_m(z) = 1, \qquad z \neq c, \tag{31}$$

and then apply Perron's reasoning.

To prove Theorem 11, we consider the following formula ([4], §5)

$$Q_{n+m}(z) = Q_n(z)Q_{m,n}(z) + k_{n+1}Q_{n-1}(z)Q_{m-1,n+1}(z).$$

Dividing on both sides by $(z - c)^{n+m}$ and letting $m \to \infty$, we get

$$Q(z) = Q_n^*(z)q_n(z) + \frac{k_{n+1}}{(z-c)^2} Q_{n-1}^*(z)q_{n+1}(z).$$

Let, now, ϵ (>0) be arbitrarily assigned. By (31) we can find a number n_ϵ such that for $n \geqq n_\epsilon$, $q_{n+1}(z)$ has no zeros outside the circle $|z - c| = \epsilon$. It follows, that for the same n (i.e., for $n \geqq n_\epsilon$) if ζ is an arbitrary zero of $Q(z)$ such that $|\zeta - c| > \epsilon$, we cannot also have $Q_n(\zeta) = 0$. For, if the latter equation were true, then it would follow that also $Q_{n-1}(\zeta) = 0$, and this is impossible in view of the relations ([4], §6)

$$P_n(z)Q_{n-1}(z) - P_{n-1}(z)Q_n(z) = (-1)^{n-1}k_1 \cdots k_n$$

and our assumption $k_i \neq 0$ $(i = 1, 2, \cdots)$.

[9] The limit function of sequence (16) is not $\equiv 0$, for

$$\lim_{|z-c|\to\infty} \frac{e^{T(z;n)}}{(z-c)^n} Q_n(z) = 1.$$

[10] An analogous theorem for (1) under (2) was first proved by Maillet [3].

By Theorem 10, it follows that the zeros of $Q(z)$ are points of proper divergence of the continued fraction (that is, at those points, the reciprocal of the continued fraction converges to zero). Furthermore, by Theorem 11, it follows that the limit points of the set of zeros of the $Q_n(z)$ $(n = 1, 2, \cdots)$ can be taken in the geometric sense, since no points of the z-plane distinct from the point c can be a zero for an infinity of $Q_n(z)$. We have, therefore,

THEOREM 12. *Under the assumptions of Theorem* 5, *the continued fraction* (3) *converges at every point of the z-plane with the exception of the limit points of the set of zeros of the polynomials* $Q_n(z)$ $(n = 1, 2, \cdots)$ *and with the possible exception of the point c. At the aforementioned limit points, the continued fraction diverges properly. In every closed domain excluding those limit points and the point c, the continued fraction converges uniformly.*

5. **Proof of Theorem 2.** By the recurrence formulas (10), we have, putting $V_n(z) = Q_n(z) \Big/ \prod_{i=1}^{n} (-c_i)$,

$$V_n(z) = \left(1 - \frac{z}{c_n}\right) V_{n-1}(z) + \frac{k_n}{c_{n-1}c_n} V_{n-2}(z).$$

Hence

$$| V_n(z) - V_{n-1}(z) | \leqq \frac{|z|}{|c_n|} | V_{n-1}(z) | + \left| \frac{k_n}{c_{n-1}c_n} \right| | V_{n-2}(z) |.$$

On the other hand, it is seen, as in the case of the derivation of (24), that

$$| V_n(z) | \leqq \prod_{i=1}^{\infty} \left(1 + \left|\frac{k_{i+1}}{c_i c_{i+1}}\right|\right) \prod_{i=1}^{\infty} \left(1 + \frac{|z|}{|c_i|}\right).$$

Recalling assumptions (5) and (6), we see that the theorem follows just as in the proof of Theorem 5.

A partial converse to Theorem 2 for a special case is given in the following remark:

If the coefficients of (3) *are real, and if*

$$\frac{k_{n+1}}{c_n c_{n+1}} > 0 \qquad (n = 1, 2, \cdots),$$

then the assumption

$$\lim_{n\to\infty} \frac{Q_n(0)}{\prod_{i=1}^{n} (-c_i)} \text{ exists} \tag{32}$$

implies (6). *Moreover, if we have*

$$k_n > 0, \qquad c_n \text{ of same sign} \qquad (n = 1, 2, \cdots),$$

then (32) *and the additional assumption*

$$\lim_{n\to\infty} \frac{Q_n'(0)}{\prod_{i=1}^{n}(-c_i)} \quad exists$$

imply both (5) *and* (6).

To see the truth of this remark we have only to observe that by (9)

$$\frac{Q_n(z)}{\prod_{i=1}^{n}(-c_i)} = \prod_{k=1}^{n}\left(1-\frac{z}{c_k}\right) + \sum_{i}^{1,n-1} \frac{k_{i+1}}{c_i c_{i+1}} \prod_{k\neq i,i+1}^{1,n}\left(1-\frac{z}{c_k}\right)$$

$$+ \sum_{i<j}^{1,n-2} \frac{k_{i+1}}{c_i c_{i+1}} \frac{k_{j+1}}{c_{j+1}c_{j+2}} \prod_{k\neq i,i+1,j+1,j+2}^{1,n}\left(1-\frac{z}{c_k}\right) + \cdots$$

and therefore

$$\frac{Q_n(0)}{\prod_{i=1}^{n}(-c_i)} = 1 + \sum_{i}^{1,n-1} \frac{k_{i+1}}{c_i c_{i+1}} + \sum_{i<j}^{1,n-2} \frac{k_{i+1}}{c_i c_{i+1}} \frac{k_{j+2}}{c_{j+1}c_{j+2}} + \cdots,$$

$$\frac{Q_n'(0)}{\prod_{i=1}^{n}(-c_i)} = -\left[\sum_{k=1}^{n}\frac{1}{c_k} + \sum_{i}^{1,n-1} \frac{k_{i+1}}{c_i c_{i+1}} \sum_{k\neq i,i+1}^{1,n} \frac{1}{c_k}\right.$$

$$\left. + \sum_{i<j}^{1,n-2} \frac{k_{i+1}}{c_i c_{i+1}} \frac{k_{j+2}}{c_{j+1}c_{j+2}} \sum_{k\neq i,i+1,j+1,j+2} \frac{1}{c_k} + \cdots\right].$$

By examining the proofs of the theorems given in §3 and that of Theorem 12, it is readily seen that we can prove by the same methods, analogous results for the continued fraction of Theorem 2.

II. Bessel's continued fraction and Lommel polynomials

6. An interesting example of a continued fraction belonging to the class discussed in §3 is provided by Bessel's continued fraction ([7], §5.6):

$$(33) \qquad \frac{1/2\nu|}{|\,z} - \frac{1/4\nu(\nu+1)\,|}{|\quad z} - \frac{1/4(\nu+1)(\nu+2)\,|}{|\quad z} - \cdots;$$

$$z \neq 0,\ \nu \text{ complex and } \neq 0, -1, -2, \cdots$$

which is "equivalent" ([4], §42) to:

$$(34) \qquad \frac{1\ |}{|\,2\nu/t} - \frac{1\quad |}{|\,2(\nu+1)/t} - \cdots, \qquad t = \frac{1}{z}.$$

The denominator of the n-th approximant of (34) is Lommel's polynomial[11]

[11] This designation is slightly inaccurate, $R_{n,\nu}(t)$ being a polynomial in $1/t$ and not in t.

$R_{n,\nu}(t)$ ([7], §9.6).[12] An important result, discovered by Hurwitz [2] and applied by him to the investigation of the zeros of Bessel's function $J_\nu(z)$, is

$$\lim_{n\to\infty} \frac{(\frac{1}{2}t)^{n+\nu}}{\Gamma(n+\nu+1)} R_{n,\nu+1}(t) = J_\nu(t) \tag{35}$$

for non-integral values of ν, the convergence being uniform in any bounded domain of the t-plane.

As an application of some of the topics discussed in part I, we shall give a simple proof of (35) with the ν restricted only as in (33): $\nu \neq 0, -1, -2, \cdots$. Other known properties of $J_\nu(z)$ will also follow from the development that will now be outlined.

The explicit expression for $J_\nu(z)$ will be needed only at one step in our discussion where we have to identify it with a derived function. Otherwise, our information is obtained directly from the functional equation ([7], §3.2)

$$J_{\nu-1}(z) + J_{\nu+1}(z) = \frac{2\nu}{z} J_\nu(z). \tag{36}$$

By (36) it is easy to verify the known fact[13] that the continued fraction (33) represents the function $J_\nu(z^{-1})/J_{\nu-1}(z^{-1}) \equiv F_\nu(z)$. Indeed, it is readily deducible from (36) that the function[14] $F_\nu(z)$ is "associated" ([4], §§61 and 67) with the continued fraction (33). But this fraction represents, by Theorem 12, a meromorphic function in z^{-1}, and it can be easily concluded, taking in consideration the uniqueness of "associatedness" ([4], §61), that this meromorphic function must be identical with $F_\nu(z)$.[15]

Consider now fraction (33) and denote its n-th approximant by $P_n(z; \nu)/Q_n(z; \nu)$. By Theorem 5, the following limits exist:

$$\lim_{n\to\infty} \frac{P_n(z;\nu)}{z^n} = p_{\nu-1}(z), \text{ say;} \qquad \lim_{n\to\infty} \frac{Q_n(z;\nu)}{z^n} = q_{\nu-1}(z), \text{ say.} \tag{37}$$

Again, by (8), it is seen that

$$P_n(z;\nu) = \frac{1}{2\nu} Q_{n-1}(z; \nu+1).$$

[12] This is not the usual definition ([7], §9.6), but it follows readily from the following properties of Lommel polynomials ([7], §9.63):

$$R_{n,\nu}(t) = \frac{2(\nu+n-1)}{t} R_{n-1,\nu}(t) - R_{n-2,\nu}(t) \qquad (n = 2, 3, \cdots),$$

$$R_{0,\nu}(t) = 1, \qquad R_{1,\nu}(t) = \frac{2\nu}{t}.$$

[13] [4], p. 299.

[14] That is, its formal expansion in powers of z^{-1}.

[15] Cf. [4], p. 353 and Theorem 20 on p. 342.

Hence

$$p_{\nu-1}(z) = \frac{1}{2\nu z} q_\nu(z)$$

and we have (by the above)

$$F_\nu(z) = \frac{1}{2\nu z} \frac{q_\nu(z)}{q_{\nu-1}(z)} \qquad (z \neq 0).$$

If we put $q_\nu(z)/J_\nu(z^{-1}) = X_\nu(z)$, the above equation becomes

$$X_\nu(z) = 2\nu z X_{\nu-1}(z),$$

a difference equation in ν having the solution

$$X_\nu(z) = \frac{1}{\psi_\nu(z)} (2z)^\nu \Gamma(\nu + 1), \qquad \psi_{\nu+1}(z) \equiv \psi_\nu(z).$$

Thus:

$$J_\nu\left(\frac{1}{z}\right) = \psi_\nu(z) \frac{1}{2^\nu \Gamma(\nu + 1) z^\nu} q_\nu(z) \qquad (z \neq 0). \tag{38}$$

It will now be shown that $\psi_\nu(z) \equiv 1$, by identifying the coefficients of the power series expansion of $q_\nu(z)$ with those of $2^\nu\Gamma(\nu + 1)z^\nu J_\nu(z^{-1})$. In fact, consider (30) with $m = 0$. It is readily seen that in our case it can be written

$$q_{\nu-1}(z) = q_\nu(z) - \frac{1}{4\nu(\nu + 1)z^2} q_{\nu+1}(z).$$

Let the expansion of $q_\nu(z)$ be (the form following by §3)

$$1 + q_1^{(\nu)} \frac{1}{z^2} + q_2^{(\nu)} \frac{1}{z^4} + \cdots.$$

Substituting in the above equation, we obtain the system of difference equations

$$q_m^{(\nu)} - q_m^{(\nu-1)} = \frac{1}{4\nu(\nu + 1)} q_{m-1}^{(\nu+1)} \qquad (m = 1, 2, \cdots),$$

and these lead very easily to our desired result.[16]

[16] Since $q_0^{(\nu)} = 1$, and since (cf. (9))

$$\frac{Q_n(z; \nu)}{z^n} = 1 + \frac{1}{z^2} \sum_{i}^{1, n-1} k_{i+1} + \Big(\sum_{i<j}^{1, n-2} k_{i+1} k_{j+2}\Big) \frac{1}{z^4} + \cdots;$$

$$k_1 = \frac{1}{2\nu}, \qquad k_i = \frac{1}{4(\nu + i - 1)(\nu + i - 2)} \qquad (i = 2, 3, \cdots),$$

taking $m = 1$ gives

$$q_1^{(\nu)} = -\frac{1}{4} \sum_{k=0}^{\infty} \frac{1}{(\nu + k)(\nu + k + 1)} = -\frac{1}{4\nu},$$

the periodic additive function vanishing by (37). Proceeding in this manner, we find (by mathematical induction)

$$q_m^{(\nu)} = \frac{(-1)^m}{4^m m!} \frac{1}{\nu(\nu + 1) \cdots (\nu + m - 1)}.$$

We have now by (38)

$$J_\nu(z) = \frac{1}{2^\nu \Gamma(\nu + 1) z^\nu} q_\nu(z).$$

This relationship permits the application of the results obtained in §3 to the function $J_\nu(z)$. Thus, Hurwitz's result (35) follows immediately from (37), since

$$R_{n,\nu}\left(\frac{1}{z}\right) = 2^n \nu(\nu + 1) \cdots (\nu + n - 1) Q_n(z; \nu) \tag{39}$$

by the definition of $R_{n,\nu}$ and Q_n ([4], §42), while the well-known fact that $z^{-\nu} J_\nu(z)$ is of genus 0 in z follows from Theorem 9.

7. For positive values of the parameter ν, the coefficients of the continued fraction (33) are seen to satisfy conditions (11) (with $c_i = 0$) and the fraction is therefore related to a moments problem ([4], §72; [6]; [1]; [5]). That means that if the power series "associated" with (33) is

$$\frac{a_0}{z} + \frac{a_1}{z^2} + \frac{a_2}{z^3} + \cdots,$$

then the infinite system of integral equations

$$\int_{-\infty}^{\infty} x^n \, d\phi(x) = a_n \qquad (n = 0, 1, \cdots) \tag{40}$$

has at least one solution $\phi(x)$ which is a bounded monotonically increasing function of the real variable x having an infinity of jump-points, the integration being taken in the Stieltjes sense. It follows further that the sequence of polynomials $Q_n(z; \nu) \equiv Q_n(z)$ considered in §6 forms for every $\nu > 0$ an orthogonal set with respect to this moments problem [5]; that is,

$$\int_{-\infty}^{\infty} Q_n(x) Q_m(x) \, d\phi(x) = 0 \qquad (n \neq m; n, m = 0, 1, \cdots),$$

where $\phi(x)$ is a solution of (40). Hence by (39), we have

$$\int_{-\infty}^{\infty} R_{n,\nu}\left(\frac{1}{x}\right) R_{m,\nu}\left(\frac{1}{x}\right) d\phi(x) = 0 \qquad (n \neq m; n, m = 0, 1, \cdots; \nu > 0). \tag{41}$$

This result is a new property of the system of Lommel polynomials.

By (41), we can obtain very simply the reality and separability properties of the zeros of Lommel's polynomials for $\nu > 0$ [5] and thence the reality of the zeros of $J_\nu(z)$ for ν real and > -1. In fact, we can apply the extensive theory of orthogonal polynomials [5] to obtain further properties of $R_{n,\nu}$. However, for a full utilization of the relations (41), it is necessary to solve the moments problem (40). This can be accomplished very easily in the present case. The fact, pointed out in §6, that the fraction (33) represents a meromorphic function in z^{-1} tells us immediately ([4], Chapter 9; [5]) that the "interval of orthogo-

nality" [5] is finite and the moments problem is "determined" ([4], §72; [5]); that is, there exists essentially only one solution. It also follows readily by "Stieltjes' inversion formula" ([6]; [4], §66)[17] that the solution $\phi(x)$ is a step-function with jump-points at the poles of the meromorphic functions in question,[18] namely, $J_\nu(z^{-1})/J_{\nu-1}(z^{-1})$ (cf. §6). Furthermore, since all $c_i = 0$ it follows [5] that the solution is "symmetric"; that is, we have with the normalization $\phi(x) = 0$

$$\phi(x) = -\phi(-x). \tag{42}$$

To solve our moments problem completely it is still necessary to determine the jumps of $\phi(x)$. Denote the positive zeros of $J_\nu(z^{-1})$ by $j_k^{(\nu)} \equiv j_k$ ($k = 1, 2, \cdots; j_{k+1} < j_k$) and the corresponding jumps of $\phi(x)$ by h_k. By (42), it follows that the jump-points of $\phi(x)$ are $\pm j_k$ ($k = 1, 2, \cdots$) and the interval of orthogonality is $(-j_1, j_1)$. Now, by Markoff's theorem ([4], §68) the continued fraction is equal (at its convergence points) to the Stieltjes integral

$$\int_{-j_1}^{j_1} \frac{d\phi(x)}{z - x}.$$

Therefore, by §6 and the form of $\phi(x)$,

$$\frac{J_\nu\left(\frac{1}{z}\right)}{J_{\nu-1}\left(\frac{1}{z}\right)} = \int_{-j_1}^{j_1} \frac{d\phi(x)}{z - x} = \sum_{k=-\infty}^{\infty} \frac{h_k}{z - j_k} \qquad (h_k = h_{-k} \text{ by } (42)),$$

the series being absolutely convergent for all $z \neq 0, \pm j_1, \pm j_2, \cdots$. It follows that h_k is equal to the residue of the function $J_\nu(z^{-1})/J_{\nu-1}(z^{-1})$ at the point j_k so that, remembering the definition of j_k,

$$h_k = \lim_{z \to j_k} \left[(z - j_k) \frac{J_\nu\left(\frac{1}{z}\right)}{J_{\nu-1}\left(\frac{1}{z}\right)} \right] = J_\nu\left(\frac{1}{j_k}\right) \frac{-j_k^2}{J'_{\nu-1}\left(\frac{1}{j_k}\right)}.$$

Hence, in view of the functional equation ([7], §3.2)

$$-J'_{\nu-1}(z) + \frac{\nu - 1}{z} J_{\nu-1}(z) = J_\nu(z),$$

we find

$$h_k = j_k^2 \qquad (k = 1, 2, \cdots).$$

The orthogonality relations (41) have thus the following form:

$$\sum_{k=-\infty}^{\infty} j_k^2 R_{n,\nu}\left(\frac{1}{j_k}\right) R_{m,\nu}\left(\frac{1}{j_k}\right) = 0 \qquad (m \neq n; m, n = 0, 1, \cdots).$$

[17] Hilfssatz 5. Cf. also A. Wintner, *Spektraltheorie der Unendlichen Matrizen*, §43.

[18] Cf. [4], p. 403, and A. Wintner, loc. cit., §45.

Bibliography

1. H. Hamburger, *Ueber eine Erweiterung des Stieltjesschen Momentenproblems*, Math. Annalen, vol. 81(1920), pp. 234–319; vol. 82(1921), pp. 120–164, 168–187.
2. A. Hurwitz, *Ueber die Nullstellen der Bessel'schen Function*, Math. Annalen, vol. 33(1889), pp. 246–266.
3. E. Maillet, *Sur les fractions continues algébriques*, Journal de l'École Polytechnique, (2), vol. 12(1908), pp. 41–62.
4. O. Perron, *Die Lehre von den Kettenbrüchen*, 2d ed., 1929.
5. J. Shohat, *Théorie Générale des Polynomes Orthogonaux de Tchebichef*, Mémorial des Sciences Mathématiques, Fasc. 66, 1934.
6. T. J. Stieltjes, *Recherches sur les Fractions Continues*, Oeuvres, vol. 2, pp. 402–566.
7. G. N. Watson, *Theory of Bessel Functions*, 1922.

University of Pennsylvania.

GENERALIZATIONS OF NON-ALTERNATING AND NON-SEPARATING TRANSFORMATIONS

By E. P. Vance

1. **Introduction.** In two papers, one by G. T. Whyburn[1] and the other by J. F. Wardwell,[2] non-alternating and non-separating transformations have been discussed. Two other transformations of the same type, completely non-alternating and completely non-separating, have also been considered but as yet have not appeared in the literature. The conditions on these four transformations are quite strong, thus limiting the applications. For example, no transformation on an acyclic space can possibly satisfy the hypothesis of completely non-separating. Also, if the transformation is non-separating, then the image is always cyclically connected. Hence it seems desirable to weaken the conditions on these four transformations, thus extending the range of possible applications. It is the purpose of this paper to present what seems to be a natural weakening or generalization of these four transformations, to investigate what theorems concerning the original stronger transformations carry over, and in most cases to consider what additional hypothesis must be added to the weakened transformations to obtain similar results.

Throughout this paper it will be assumed that the space A is a compact metric continuum. All transformations $T(A) = B$ which are considered are assumed to be single valued and continuous.

Six new transformations are defined in this paper. A transformation T is said to be *weakly non-alternating*[3] if, for any two distinct points x and y of B, $T^{-1}(x)$ does not separate two points of $T^{-1}(y)$ non-degenerately in A.[4] A transformation T is said to be *weakly non-separating*[5] if, for any point x of B, $T^{-1}(x)$ does not separate any two points of A non-degenerately. A transformation T is said to be *weakly completely non-alternating*[6] if, for any two points x and y

Received June 29, 1939; presented to the American Mathematical Society, April, 1939; abstract, Bulletin of the American Mathematical Society, vol. 45(1939), p. 237. The writer wishes to express his appreciation for the help given by Professor W. L. Ayres, who suggested the investigation of this problem and gave valuable criticism and advice during the preparation of this paper.

[1] G. T. Whyburn, *Non-alternating transformations*, American Journal of Mathematics, vol. 56(1934), pp. 294–302.

[2] J. F. Wardwell, *Non-separating transformations*, this Journal, vol. 2(1936), pp. 745–750.

[3] A transformation T is said to be *non-alternating* if, for any two distinct points x and y of B, $T^{-1}(x)$ does not separate $T^{-1}(y)$ in A. G. T. Whyburn, loc. cit., p. 294.

[4] A subset K will be said to *separate* a subset H *non-degenerately* if K separates H in A and no single point of K separates H in A.

[5] A transformation T is said to be *non-separating* if, for any point x of B, $T^{-1}(x)$ does not separate A. J. F. Wardwell, loc. cit., p. 745.

[6] A transformation T is said to be *completely non-alternating* if, for any two points x and y of B and any closed subset K of $T^{-1}(x)$, K does not separate $T^{-1}(y)$ in A.

of B and any closed subset K of $T^{-1}(x)$, K does not separate two points of $T^{-1}(y)$ non-degenerately in A. (x may be identical with y.) A transformation T is said to be *weakly completely non-separating*[7] if, for any point x of B and any closed subset K of $T^{-1}(x)$, K does not separate two points of A non-degenerately. Two more transformations will be defined later, locally non-alternating and locally non-separating.

In considering the relations between these four transformations and the four originally defined, the transformations previously discussed imply the corresponding weaker ones in all cases. Among the transformations themselves, the completely non-separating type implies both the completely non-alternating and the non-separating type, both of which in turn imply the non-alternating. These implications follow directly from the definitions. By simple examples it is easily seen that there exist no other relations between these eight transformations without added conditions.

The following example shows the difference between the new and the old definitions. Let A be defined as the closed rectangle $-2 \leqq x \leqq -1, 0 \leqq y \leqq 1$ plus the closed rectangle $1 \leqq x \leqq 2$, $0 \leqq y \leqq 1$ plus the closed line interval $-1 \leqq x \leqq 1$, $y = 0$. Let T be defined as the following two steps: (1) a folding of the negative half of A into the positive half and (2) then identifying the line $0 \leqq x \leqq 1$, $y = 0$ with the point (1, 0). This transformation satisfies the conditions of a weakly completely non-separating transformation and therefore the conditions of all the other weaker transformations; and yet it is not non-alternating and therefore is none of the stronger transformations.

Since any continuous transformation of A into B is equivalent to an upper semi-continuous decomposition[8] of A into disjoint closed sets where the hyperspace of the decomposition is homeomorphic with B, any one of these transformations defined is equivalent to an upper semi-continuous decomposition of A into sets which have the required properties stated in the particular definition.

2. Characteristic properties.

THEOREM 2.1. *In order that T be weakly non-alternating it is necessary and sufficient that, for any element x of B,*

$T^{-1}(x)$ separates two points a_1 and a_2 in A

implies either

(1) *x separates $T(a_1)$ and $T(a_2)$ in B; or*

(2) *for any separation $A - T^{-1}(x) = A_1 + A_2$ and any pair of points u of A_1 and v of A_2 such that $T(u) = T(v)$, u and v are separated in A by a single point of $T^{-1}(x)$.*

Sufficiency. Suppose T is not weakly non-alternating. Then some set $T^{-1}(x)$ separates two points a_1 and a_2 of some $T^{-1}(y)$; that is, $A - T^{-1}(x) = A_1 + A_2$,

[7] A transformation T is said to be *completely non-separating* if, for any point x of B and any closed subset K of $T^{-1}(x)$, K does not separate A.

[8] R. L. Moore, *Concerning upper semi-continuous collections of continua*, Transactions of the American Mathematical Society, vol. 27(1925), pp. 416–428.

where A_1 contains a_1, A_2 contains a_2 and $\bar{A}_1 \cdot A_2 = A_1 \cdot \bar{A}_2 = 0$. Since a point x cannot separate a single point y of B, alternative (1) is impossible and alternative (2) must hold. But by (2) with $u = a_1$ and $v = a_2$ the separation above fails to be non-degenerate and the sufficiency is proved.

Necessity. Since $T^{-1}(x)$ separates a_1 and a_2 in A, $A - T^{-1}(x) = A_1 + A_2$, where a_i lies in A_i. Suppose no set $T^{-1}(z)$, z a point of B, has points in both A_1 and A_2. Let $B_1 = T(A_1)$ and $B_2 = T(A_2)$. Then $B_1 \cdot B_2 = 0$. Also $B_1 \cdot \bar{B}_2 = 0$. For suppose b is a point of $B_1 \cdot \bar{B}_2$; then[9] $T^{-1}(b) \cdot \overline{T^{-1}(B_2)} \neq 0$ and lies in $A_1 \cdot \bar{A}_2$, contrary to the assumption that A_1 and A_2 are mutually separated. Similarly $\bar{B}_1 \cdot B_2 = 0$. Thus $B - x = B_1 + B_2$ is a separation; that is, x separates $T(A_1)$ and $T(A_2)$ in B, and alternative (1) is obtained. Suppose there is a set $T^{-1}(z)$ such that $T^{-1}(z) \cdot A_1 \neq 0 \neq T^{-1}(z) \cdot A_2$. Let u and v be points of these two sets. Since T is weakly non-alternating, it follows that u and v may be separated in A by a single point of $T^{-1}(x)$. Further, since any two points u, a point of A_1, and v, a point of A_2, such that $T(u) = T(v)$, belong to such a set $T^{-1}(z)$, alternative (2) has been proved.

THEOREM 2.2. *If T is weakly non-separating, then for any cut point b of B, $T^{-1}(b)$ contains a cut point of A.*

If b is a cut point of B, that is, $B - b = B_1 + B_2$ is a separation, then by continuity of the transformation $A - T^{-1}(b) = T^{-1}(B_1) + T^{-1}(B_2)$ is a separation. Since T is weakly non-separating, $T^{-1}(b)$ does not separate any two points u and v unless a single point does. Therefore, since $T^{-1}(b)$ separates any point u in $T^{-1}(B_1)$ and any point v in $T^{-1}(B_2)$, a single point would separate; that is, $T^{-1}(b)$ contains a cut point of A.

The above condition is necessary as we have seen but not sufficient, even though we assume in addition that T is weakly non-alternating.

THEOREM 2.3. *If T is weakly completely non-alternating, then every open set of each set $T^{-1}(x)$ contains a cut point of A.*

Assume $T^{-1}(x)$ contains an open set U which contains no cut point of A. For every p which is a point of U, there exists a neighborhood U_p such that $\bar{U}_p$ lies in U. Therefore the $B(U_p)$ separates U into U_p and $U - \bar{U}_p$. If we consider the $B(U_p)$ as the closed set K in the definition of a weakly completely non-alternating transformation, we have $B(U_p)$ a closed set lying in $T^{-1}(x)$ and separating a point $u = p$ of $T^{-1}(x)$ in $\bar{U}_p$ and a point v of $T^{-1}(x)$ in $U - \bar{U}_p$. Now since U contains no cut points, $B(U_p)$ contains no cut points; that is, no single point of $B(U_p)$ separates u and v. This is contrary to the fact that T is weakly completely non-alternating.

COROLLARY 2.31. *If T is weakly completely non-separating, then every open set of each set $T^{-1}(x)$ contains a cut point of A.*

Again the above conditions are necessary but not sufficient, even though T is assumed to be weakly non-separating, which implies weakly non-alternating.

[9] G. T. Whyburn, *Non-alternating transformations*, loc. cit.

3. **Applications to Peano space.** In order to characterize further these transformations it will be supposed in this section that A, and therefore B, will be locally connected continua.

THEOREM 3.1. *If T is non-alternating and, for every cut point b of B, $T^{-1}(b)$ contains no non-degenerate irreducible cut set between two points of A, then T is weakly non-separating.*

If T is not weakly non-separating, there exists a separation between two points u and v, $A - T^{-1}(x) = A_1 + A_2$, where A_1 contains u and A_2 contains v, such that no single point of $T^{-1}(x)$ separates u and v. Since T is non-alternating, there exists no $T^{-1}(y)$ such that the set $T^{-1}(y)$ contains points in both A_1 and A_2. Therefore A_1 is the sum of certain sets $T^{-1}(y)$ and A_2 is the sum of other sets $T^{-1}(y)$. Thus because of the continuity, x separates the sets $T(A_1)$ and $T(A_2)$; that is, x is a cut point of B. Therefore by hypothesis $T^{-1}(x)$ contains no non-degenerate irreducible cut set between any two points. Since in a Peano space every cut set between two points contains an irreducible cut set between these two points,[10] this cut set can be reduced to a point; that is, the above separation between u and v can be made by one point of $T^{-1}(x)$, and this is contrary to the assumption.

An example shows that the above condition is not necessary for a weakly non-separating transformation.

THEOREM 3.2. *If T is completely non-alternating and, for every cut point b of B, $T^{-1}(b)$ contains no non-degenerate irreducible cut set between two points of A, then T is weakly completely non-separating.*

Assume T is not weakly completely non-separating; that is, there exist a closed set K in $T^{-1}(x)$ and two points u and v such that $A - K = A_1 + A_2$ is a separation, where A_1 contains u and A_2 contains v and no single point of K separates u and v. Since T is completely non-alternating, no closed set K in $T^{-1}(x)$ separates any $T^{-1}(y)$ (y ranges over B including x). Then each set $T^{-1}(y)$ lies entirely in one of the sets A_i, say A_1.

Suppose that A_1 lies entirely in $T^{-1}(x) - K$. Consider any point s of A_1. Let d equal $\rho(s, K)$, the minimum distance between s and the closed set K. Define K_1 to be all points $\{p\}$ such that p is an element of A_1 and such that $\rho(p, K) = \frac{1}{2}d$. Thus K_1 is a closed subset of $T^{-1}(x)$ such that $A - K_1 = A_1^1 + A_2^1$ is a separation, where $A_1^1 =$ all points $\{p'\}$ such that p' is a point of A_1 and $\rho(p', K) > \frac{1}{2}d$, and where $A_2^1 = A - (A_1^1 + K_1)$. But since A_1^1 contains points of $T^{-1}(x)$, for example s, and A_2^1 contains points of $T^{-1}(x)$, namely, those in K, the fact that T is completely non-alternating is contradicted. Thus $A_1 - T^{-1}(x) \neq 0$. Then A_2 is the sum of certain sets $T^{-1}(y)$, A_1 is $T^{-1}(x) - K$ plus certain other sets $T^{-1}(y)$, and there is at least one $T^{-1}(y)$ in each set. Therefore by continuity $B - x = T(A_1 - T^{-1}(x)) + T(A_2)$ is a separation; that is, x is a cut point of B. Therefore by hypothesis $T^{-1}(x)$ con-

[10] G. T. Whyburn, *Concerning irreducible cuttings of continua*, Fundamenta Mathematicae, vol. 13(1928), pp. 47–57.

tains no non-degenerate irreducible cut set between any two points. But in Peano space each such cut set contains an irreducible cut set which in this case must be degenerate.[11] Hence K contains a point which separates u and v in A, contrary to the assumption. Therefore T is weakly completely non-separating.

THEOREM 3.3. *If T is weakly completely non-separating, then, for any cut point b of B, $T^{-1}(b)$ contains no non-degenerate irreducible cut set between any two points of A.*

Consider any cut point b of B, and let $B - b = B_1 + B_2$ be a separation. Thus because of the continuity $T^{-1}(B) - T^{-1}(b) = T^{-1}(B_1) + T^{-1}(B_2)$ or $A - T^{-1}(b) = T^{-1}(B_1) + T^{-1}(B_2)$. Since T is weakly completely non-separating, for any closed set K in $T^{-1}(b)$, K does not separate A between two points unless a single point does. If K does not separate A, then K is not a cut set. If K separates A between two points, then a single point of K will separate; that is, every cut set of $T^{-1}(b)$ contains no non-degenerate irreducible cut set.

Thus the irreducible cut set condition is also seen to be necessary even in non-Peano space. It is possible to set up examples to show that the necessity on Theorems 3.1 and 3.2 is not true even though only the weakened types of non-alternating transformations are assumed.

It has been noticed that these theorems (3.1 and 3.2) have been proved only in Peano space. The following example will indicate clearly why they are not possible in non-Peano space. Let A be defined as follows: Consider a non-dense perfect set on the closed interval $0 \leqq x \leqq 1$. Join all points x of the non-dense perfect set by straight line intervals L_x to the point $(\frac{1}{2}, 2)$. Also consider the sequence of points in the line L_0 from $(\frac{1}{2}, 2)$ to 0 whose y coördinates are $1, \frac{3}{2}, \frac{5}{4}, \frac{9}{8}, \frac{17}{16}, \cdots$. This sequence will be called $\{a_i\}$. In a similar fashion form $\{b_i\}$, points in L_1 with respective y coördinates. Define R_i as the closed interval joining a_i to b_i for all i. Let the set A be defined as the sum of the sets L_x for all x of the perfect set and the sum of the intervals R_i. Let T be defined as a homeomorphism for all points of A, with one exception. Let all the points on the line R_i where $y = 1$ be carried into one point p. The transformation T is clearly completely non-alternating; for all points except p have a single point as the inverse set and therefore cannot be separated. No single point separates $T^{-1}(p)$. Therefore the only remaining possibility is that a closed subset K of $T^{-1}(p)$ may separate two points of $T^{-1}(p)$. This is clearly impossible since any two points of $T^{-1}(p) - K$ can be joined in A by an arc outside K. Moreover, the only cut point of B is p, and no matter what two points a closed subset K of $T^{-1}(p)$ separates, the cut set can always be reduced, so $T^{-1}(p)$ contains no irreducible cut set, degenerate or non-degenerate, between any two points of A. Yet T is not even weakly non-separating; for $T^{-1}(p)$ separates the point (0, 0) from the point $(\frac{1}{2}, 2)$, and no single point of $T^{-1}(p)$ does.

[11] G. T. Whyburn, *Concerning irreducible cuttings of continua*, loc. cit., Theorem 8.

THEOREM 3.4. *If A is a cyclically connected Peano space,* (1) *any weakly non-alternating transformation is a non-alternating transformation,* (2) *any weakly non-separating transformation is a non-separating transformation,* (3) *any weakly completely non-alternating transformation is a completely non-alternating transformation, and* (4) *any weakly completely non-separating transformation is completely non-separating.*

For the proof of these statements it is sufficient to note that since A is cyclic, no single point of A separates any two points of A.

4. **Product and factor theorems.** In this section the spaces A, B, and C referred to are compact metric continua. Let $T_1(A) = B$ and $T_2(B) = C$, and let $T = T_2T_1$. In other words T is the result of first applying T_1 to A and then applying T_2 to $T_1(A)$, so that we have $T(A) = T_2(T_1(A)) = T_2(B) = C$.

THEOREM 4.1. *If T_1 is monotone and non-separating, and T_2 is weakly non-separating, then T is non-separating.*

Assume T is not non-separating; that is, there exists a separation $A - T^{-1}(x) = A_1 + A_2$. Consider $T_1(A) - T_1(T^{-1}(x)) = T_1(A_1) + T_1(A_2)$. The set $T_1(A_1) \cdot T_1(A_2)$ is vacuous; for if not, consider p a point of $T_1(A_1) \cdot T_1(A_2)$. The set $T_1^{-1}(p)$ is a subset of $A - T^{-1}(x)$, has points in both A_1 and A_2, and is connected since T_1 is monotone; and this contradicts the fact that A_1 and A_2 were separated. Therefore $T_1(A) - T_1(T^{-1}(x)) = T_1(A_1) + T_1(A_2)$ is a separation.[12] That is, $B - T_2^{-1}(x) = T_1(A_1) + T_1(A_2)$ is a separation. Suppose a single point p of $T_2^{-1}(x)$ would separate; that is, $B - p = K_1 + K_2$. Then by continuity $T_1^{-1}(B) - T_1^{-1}(p) = T_1^{-1}(K_1) + T_1^{-1}(K_2)$ would still be a separation, that is, $T_1^{-1}(p)$ would separate A, and the fact that T_1 is non-separating is contradicted. Therefore, no single point of $T_2^{-1}(x)$ in the above separation does the separating. But this fact contradicts the fact that T_2 is weakly non-separating and therefore our assumption is false. Thus T is non-separating.

THEOREM 4.2. *If T_1 is monotone and non-separating and T_2 is weakly non-alternating, then T is non-alternating.*

Assume T is not non-alternating, that is, there exists a separation

$$A - T^{-1}(x) = A_1 + A_2, \tag{1}$$

where some set $T^{-1}(y)$ has points in both A_1 and A_2. Consider

$$T_1(A) - T_1T^{-1}(x) = T_1(A_1) + T_1(A_2). \tag{2}$$

As in the proof of Theorem 4.1, (2) is still a separation because T_1 is monotone. Since $A_i \cdot T^{-1}(y) \neq 0$, $T_1(A_i) \cdot T_1(T^{-1}(y)) \neq 0$; that is, $T_1(A_i) \cdot T_2^{-1}(y) \neq 0$, where $i = 1$ and 2. Therefore

$$B - T_2^{-1}(x) = T_1(A_1) + T_1(A_2), \tag{3}$$

[12] G. T. Whyburn, *Non-alternating transformations*, loc. cit., p. 296, Theorem 2.1.

where both $T_1(A_1)$ and $T_1(A_2)$ have points in common with the set $T_2^{-1}(y)$. Suppose a single point of $T_2^{-1}(x)$ separated. Again as in the proof of Theorem 4.1 this contradicts the fact that T_1 is non-separating. Therefore (3) is a separation and no single point of $T_2^{-1}(x)$ does the separating. This contradicts the fact that T_2 is weakly non-alternating. Thus T is non-alternating.

THEOREM 4.3. *If T is weakly non-separating and T_1 is non-alternating, then T_2 must be weakly non-separating.*

If T_2 were not weakly non-separating, there would exist some point x of C and two points u and v of B such that $B - T_2^{-1}(x) = B_1 + B_2$ would be a separation of B between u and v, and such that no single point of $T_2^{-1}(x)$ would separate u and v in B. Since T_1 is continuous, $T_1^{-1}(B) - T_1^{-1}(T_2^{-1}(x)) = T_1^{-1}(B_1) + T_1^{-1}(B_2)$ or

$$A - T^{-1}(x) = T_1^{-1}(B_1) + T_1^{-1}(B_2) \tag{1}$$

is still a separation between $T_1^{-1}(u)$ and $T_1^{-1}(v)$. But T is weakly non-separating. Therefore if u_1 is a point of $T_1^{-1}(u)$ and v_1 is a point of $T_1^{-1}(v)$, there is a point p of the set $T^{-1}(x)$ such that

$$A - p = A_1 + A_2 \tag{2}$$

is a separation and A_1 contains u_1 and A_2 contains v_1. Consider

$$B - T_1(p) = T_1\{A_1 - T_1^{-1}(T_1(p))\} + T_1\{A_2 - T_1^{-1}(T_1(p))\}. \tag{3}$$

Neither of these two sets on the right is vacuous, for the first contains $T_1(u_1)$ and the second contains $T_1(v_1)$ since $T_1^{-1}(T_1(p))$ is a subset of $T^{-1}(x)$. Thus (3) is a separation of B. For suppose the two sets have a point y in common. The point y cannot be $T_1(p)$. Then

$$\begin{aligned} T_1^{-1}(B) - T_1^{-1}(T_1(p)) &= A - T_1^{-1}(T_1(p)) \\ &= \{A_1 - T_1^{-1}(T_1(p))\} + \{A_2 - T_1^{-1}(T_1(p))\} \end{aligned} \tag{4}$$

is a separation of A by (2) and both sets on the right contain points of $T_1^{-1}(y)$. Therefore $T_1^{-1}(y)$ would be separated by $T_1^{-1}(T_1(p))$, and this is impossible since T_1 is non-alternating. Thus we have shown that (3) is a separation of B by a single point of $T^{-1}(x)$ between u and v and the assumption is contradicted. Therefore T_2 is weakly non-separating.

THEOREM 4.4. *If T is weakly non-alternating and T_1 is non-alternating, then T_2 must be weakly non-alternating.*

If T_2 were not weakly non-alternating, there would exist two points x and y of C such that $B - T_2^{-1}(x) = B_1 + B_2$ would be a separation, where $B_1 \cdot T_2^{-1}(y)$ contains a point u and $B_2 \cdot T_2^{-1}(y)$ contains a point v, and such that no single point of $T_2^{-1}(x)$ would separate u and v in B. Since T_1 is continuous, $T_1^{-1}(B) - T_1^{-1}(T_2^{-1}(x)) = T_1^{-1}(B_1) + T_1^{-1}(B_2)$ or

$$A - T^{-1}(x) = T_1^{-1}(B_1) + T_1^{-1}(B_2) \tag{1}$$

would still be a separation between $T_1^{-1}(u)$ and $T_1^{-1}(v)$, that is, between points of $T^{-1}(y)$. But T is weakly non-alternating. Therefore

$$(2) \qquad A - p = A_1 + A_2$$

is a separation, where p is a point of the set $T^{-1}(x)$ and A_1 contains u_1, a point of $T^{-1}(y)$, and A_2 contains v_1, a point of $T^{-1}(y)$. Consider

$$(3) \qquad B - T_1(p) = T_1\{A_1 - T_1^{-1}(T_1(p))\} + T_1\{A_2 - T_1^{-1}(T_1(p))\}.$$

Again neither of the two sets on the right is vacuous, for the first contains $T_1(u_1)$ and the second contains $T_1(v_1)$. Thus (3) is a separation of B. For suppose the two sets have a point q in common. The point q is not equal to $T_1(p)$. Then

$$(4) \qquad \begin{aligned} T_1^{-1}(B) - T_1^{-1}(T_1(p)) &= A - T_1^{-1}(T_1(p)) \\ &= \{A_1 - T_1^{-1}(T_1(p))\} + \{A_2 - T_1^{-1}(T_1(p))\} \end{aligned}$$

is a separation of A by (2) and both sets on the right contain points of $T_1^{-1}(q)$. Therefore $T_1^{-1}(q)$ would be separated by $T_1^{-1}(T_1(p))$ and this is impossible since T_1 is non-alternating. Thus a single point $T_1(p)$ of $T_2^{-1}(x)$ would separate B between u and v. But this is contrary to our assumption. Thus T_2 is weakly non-alternating.

THEOREM 4.5. *If T is weakly non-separating and T_1 is non-separating, then T_2 must be non-separating.*

If T_2 were not non-separating, there would exist a separation $B - T_2^{-1}(x) = B_1 + B_2$. Consider $T_1^{-1}(B) - T_1^{-1}(T_2^{-1}(x)) = T_1^{-1}(B_1) + T_1^{-1}(B_2)$ or

$$(1) \qquad A - T^{-1}(x) = T_1^{-1}(B_1) + T_1^{-1}(B_2).$$

This is still a separation because of the continuity of T_1. Therefore since T is weakly non-separating and there is a separation of A by the set $T^{-1}(x)$, a single point p of $T^{-1}(x)$ must separate. Therefore $A - p = A_1 + A_2$, where A_i contains a point u_i of $T_1^{-1}(B_i)$ for i equal 1 and 2. Since p is a point of A, $T_1(p)$ is a point of B; and since T_1 is non-separating, $A - T_1^{-1}(T_1(p))$ is a connected set. Also $T_1^{-1}(T_1(p))$ lies in $T^{-1}(x)$ since p is a point of $T^{-1}(x)$. Therefore $A - T_1^{-1}(T_1(p))$ contains $A - T^{-1}(x)$. Thus $A - p = A_1 + A_2$ contains the connected set $A - T_1^{-1}(T_1(p))$ which contains $A - T^{-1}(x)$. Therefore A_1 or A_2, say A_1, contains $A - T^{-1}(x)$. It follows that A_2 is contained in the set $T^{-1}(x) - p$; that is, $T^{-1}(x)$ contains A_2. But A_2 contains a point u_2 of $T_1^{-1}(B_2)$ which lies in $A - T^{-1}(x)$. Hence we have the contradiction that $T^{-1}(x)$ contains a point which does not belong to $T^{-1}(x)$. Therefore T_2 is non-separating.

THEOREM 4.6. *If T is weakly non-alternating and T_1 is non-separating, then T_2 must be non-alternating.*

Assume T_2 is not non-alternating; that is, there exists a separation $B - T_2^{-1}(x) = B_1 + B_2$ such that some set $T_2^{-1}(y)$ has points in both B_1 and B_2. Consider $T_1^{-1}(B) - T_1^{-1}(T_2^{-1}(x)) = T_1^{-1}(B_1) + T_1^{-1}(B_2)$ or

$$(1) \qquad A - T^{-1}(x) = T_1^{-1}(B_1) + T_1^{-1}(B_2).$$

This is still a separation because of the continuity of the transformation. Also from the fact that $B_i \cdot T_2^{-1}(y) \neq 0$, $T_1^{-1}(B_i) \cdot T_1^{-1}(T_2^{-1}(y)) \neq 0$; that is, $T_1^{-1}(B_i) \cdot T^{-1}(y) \neq 0$ with i equal 1 and 2. Therefore in (1) we have $T^{-1}(x)$ separating two points u and v of $T^{-1}(y)$. But since T is weakly non-alternating, a single point of the set $T^{-1}(x)$ must separate u and v; that is, $A - q = A_1 + A_2$, where A_1 contains u and A_2 contains v. Since q is a point of A, $T_1(q)$ is a point of B, and since T_1 is non-separating, $A - T_1^{-1}(T_1(q))$ is connected. Now $T_1^{-1}(T_1(q))$ is contained in $T^{-1}(x)$ since q is a point of $T^{-1}(x)$. Therefore $A - T^{-1}(T_1(q))$ contains $A - T^{-1}(x)$ and $A - q$ contains $A - T_1^{-1}(T_1(q))$. Thus we have $A - q = A_1 + A_2$ containing the connected set $A - T_1^{-1}(T_1(q))$ which contains $A - T^{-1}(x)$. Therefore the set $A - T^{-1}(x)$ must lie in A_1 or A_2, say A_1. Hence A_2 is contained in $T^{-1}(x) - q$; that is, A_2 lies in $T^{-1}(x)$. But since A_2 had at least one point in common with the set $T^{-1}(y)$, there is at least one point of $T^{-1}(y)$ in $T^{-1}(x)$, and this is impossible. Thus T_2 is non-alternating.

By simple examples it is seen that these are the only possible theorems. For even though T_1 is non-alternating and monotone and T_2 weakly non-separating, T may fail to be weakly non-alternating. It is possible to have T weakly non-separating and T_1 weakly non-separating, and yet T_2 will not necessarily be weakly non-alternating. Also it is possible to have T weakly non-separating and T_1 even non-alternating and yet it is not necessary for T_2 to be non-alternating.

5. **Applications to special curves and surfaces.** In applying these transformations to the special curves and surfaces considered by Whyburn and Wardwell, there are no new applications; in the case of the boundary curve nothing can be said unless other conditions are assumed; in the case of the dendrite, any continuous transformation on a dendrite is weakly completely non-separating; and in the case of the simple closed curve and topological sphere, these weaker transformations reduce to those already considered.

6. **Locally non-separating transformations.** A continuous transformation will be called *locally non-separating at a point* provided for any point p of A and any neighborhood U_p of p, there exists a neighborhood V_p, lying in U_p such that if u and v are two points of $V_p - T^{-1}(T(p))$, then $u + v$ lies in a connected subset of $U_p - T^{-1}(T(p))$. A continuous transformation will be called *locally non-separating on a space* A if it is locally non-separating at every point of A.

It is to be noted by examples that a non-separating transformation is not necessarily a locally non-separating transformation, and conversely.

THEOREM 6.1. *If T is locally non-separating and for no point x of B does $T^{-1}(x)$ contain an open set, then T is completely non-separating (and therefore non-separating).*

If T is not non-separating, there exists a separation $A - T^{-1}(x) = A_1 + A_2$. Since $T^{-1}(x)$ contains no open set for any x, it is possible to let $T^{-1}(x) = X_1 + X_2$, where X_1 consists of those points that are limit points of A_1 and X_2 those

points that are limit points of A_2. If there were a point which was a limit point of both A_1 and A_2, T would not be locally non-separating. Thus $X_1 \cdot X_2 = 0$. Similarly, a point of $X_1 \cdot \bar{X}_2$ or $\bar{X}_1 \cdot X_2$ would be a limit point of both A_1 and A_2; hence $X_1 \cdot \bar{X}_2 = \bar{X}_1 \cdot X_2 = 0$. Also neither set A_i contains a limit point of X_j. For X_j lies in the set $T^{-1}(x)$ which is closed. Therefore $A = (A_1 + X_1) + (A_2 + X_2)$ is a separation of A, which fact contradicts A being connected. Therefore T is non-separating.

Since T is non-separating, for any x, $A - T^{-1}(x)$ is a connected set. If T were not completely non-separating, there would exist some closed subset K in some $T^{-1}(x)$ such that $A - K = A_1 + A_2$ would be a separation. Therefore $A - T^{-1}(x)$ being connected must lie wholly in A_1 or wholly in A_2. Suppose it lies in A_1. Then A_2, an open set, would be contained in $T^{-1}(x)$, and this is impossible by hypothesis. Therefore T is completely non-separating.

THEOREM 6.2. *In order for T to be locally non-separating it is necessary and sufficient that for any point p of A and any neighborhood U_p there exist a neighborhood W_p such that all the components of $W_p - T^{-1}(T(p))$ belong to a single component of $U_p - T^{-1}(T(p))$.*

Necessity. Let p be a point of A and U_p any neighborhood of p. Since T is locally non-separating, there exists a neighborhood W_p such that any two points u and v of $W_p - T^{-1}(T(p))$ lie in a connected subset of $U_p - T^{-1}(T(p))$. Let u remain fixed and v range over all points of $W_p - T^{-1}(T(p))$. Then each such point v lies with u in a connected subset C_v of $U_p - T^{-1}(T(p))$. Since each two sets have the same point u in common, $\sum C_v$ is connected and thus lies in a component of $U_p - T^{-1}(T(p))$.

Sufficiency. Since all the components of $W_p - T^{-1}(T(p))$ belong to one component of $V_p - T^{-1}(T(p))$, any two points of $W_p - T^{-1}(T(p))$ will lie in a connected set in $V_p - T^{-1}(T(p))$.

THEOREM 6.3. *If T is locally non-separating at a point p, then for any U_p at most one component of $U_p - T^{-1}(T(p))$ has p as a limit point.*

Consider any neighborhood U_p of p, and the corresponding set $U_p - T^{-1}(T(p))$. If $U_p - T^{-1}(T(p))$ is connected, the theorem is proved. Therefore assume it not connected; that is, $U_p - T^{-1}(T(p)) = S_1 + S_2$ is a separation. Also assume there are at least two components which have p as a limit point; that is, $U_p - T^{-1}(T(p)) = C_1 + C_2 + \cdots$. Now since T is locally non-separating at p, there exists a W_p lying in U_p such that all the components of $W_p - T^{-1}(T(p))$ lie in one component of $U_p - T^{-1}(T(p))$. But $C_1 \cdot W_p \neq 0$ and $C_2 \cdot W_p \neq 0$ since both sets C_i have p as a limit point. But these are subsets of two different components of $U_p - T^{-1}(T(p))$, and the fact that T is locally non-separating is contradicted. Thus at most one component of $U_p - T^{-1}(T(p))$ has p as a limit point.

THEOREM 6.4. *If T is locally non-separating, then for any point p of A and for any sufficiently small neighborhood U_p there exists a neighborhood W_p lying in U_p such that, if C_p is the component of U_p containing p, then*

(1) *if A is locally connected at p, then $W_p - C_p = 0$; or*

(2) *if A is not locally connected at p, then $T(W_p - C_p) = T(p)$.*

We first prove (1). Since A is locally connected at p, for any neighborhood U_p of p, there exists a neighborhood W_p lying in U_p such that any point of W_p can be joined to p by a connected set in U_p; that is, any point of W_p lies in C_p. Therefore $W_p - C_p = 0$.

Next consider (2). Since A is not locally connected at p, there exists a neighborhood U' of p such that any neighborhood V' of p contained in U' contains a point q which does not lie with p in any connected subset of U'; that is, p and q belong to distinct components of U'. This is obviously still true when U' is replaced by any neighborhood U_p contained in it.

Since T is locally non-separating, by Theorem 6.2, there exists a neighborhood W'_p such that all components of $W'_p - T^{-1}(T(p))$ belong to a single component of $U_p - T^{-1}(T(p))$. If this is the component C_p, let $W_p = W'_p$. If this is some other component C of U_p, there is a neighborhood W_p which contains no point of C since p is not a limit point of C. Then $W_p - C_p - T^{-1}(T(p)) = 0$ in both cases. Hence $W_p - C_p = 0$, or $T(W_p - C_p) = T(p)$. But every neighborhood W_p lying in U_p contains a point q of the previously mentioned q's, so $W_p - C_p \neq 0$ and we have $T(W_p - C_p) = T(p)$.

THEOREM 6.5. *If T is locally non-separating, then B is a Peano space.*

If A is locally connected at all points, so is B by continuity of T. Suppose A is not locally connected at some point q where $T(q) = p$. Consider any neighborhood U_p of p. Since T is continuous, there exists a neighborhood R_q such that $T(R_q)$ lies in U_p. Since A is not locally connected at q, by Theorem 6.4 there exists a V_q such that V_q lies in R_q and $T(V_q - C_q) = T(q) = p$, where C_q is the component of R_q containing q. For each q where $T(q) = p$ there exists such a V_q. (If A is locally connected at a point q, $V_q - C_q = 0$ by Theorem 6.4.) Therefore $\sum V_q$ covers $T^{-1}(p)$. By the Borel Theorem there exists a finite number of the sets V_q which cover $T^{-1}(p)$, namely, $\sum_{i=1}^{k} V_{q_i} = G$. Then $T(G)$ lies in U_p and it follows that $T(G)$ contains some open set W_p about p lying in U_p. For suppose it does not. If not, for each n there exists an x_n such that $\rho(x_n, p) \leqq n^{-1}$, where x_n is not a point of $T(G)$. Therefore $T^{-1}(x_n)$ is not contained in G. Let Y_n be a point of $T^{-1}(x_n)$. Since A is compact, $\sum y_n$ contains a subsequence $\sum y_{n_i}$ which converges to a point x. As G is open and no point y_n belongs to G, x is not in G. By continuity $T(y_{n_i}) = x_{n_i}$ approaches $T(x)$, which is equal to p because $\rho(x_{n_i}, p) \leqq n_i^{-1}$. Therefore x is a point of $T^{-1}(p)$, that is, x is one of the q's mentioned above; but x was not covered by G, and thus a contradiction is reached.

Now, for each q_i, $V_{q_i} - C_{q_i} = 0$ or $T(V_{q_i} - C_{q_i}) = T(q_i) = p$. That is, under the transformation all the V_{q_i}'s except C_{q_i} go into the same point as q_i, namely, p; and $T(C_{q_i})$ lies in U_p. Thus U_p contains $T(G)$ contains W_p, and U_p contains $T(\sum_{i=1}^{k} C_{q_i})$ contains W_p. But $T(\sum_{i=1}^{k} C_{q_i})$ is a finite number of connected sets. Thus W_p which lies in U_p is made up of at most a finite number of

components. Let H be the component of W_p containing p. As $W_p - H$ is a finite number of components, there is a neighborhood Y_p containing no points of $W_p - H$. Thus any neighborhood U_p contains a neighborhood Y_p, every point of which lies with p in a connected subset of U_p, namely, H. Therefore B is locally connected at p.

Examples will show that T may be locally non-separating and yet B may contain a cut point.

THEOREM 6.6. *If T is locally non-separating and for no x, a point of B, does $T^{-1}(x)$ contain an open set, then B has no cut points.*

Since T is locally non-separating and for no x, a point of B, does $T^{-1}(x)$ contain an open set, by Theorem 6.1 T is non-separating. Since T is non-separating, B contains no cut points by Theorem 2.1 of Wardwell's paper.

THEOREM 6.7. *If A is locally connected and T is locally non-separating, and for no point x of B does $T^{-1}(x)$ contain an open set, then*

(1) *there exists a unique true cyclic element E_a of A such that $T(E_a) = B$;*

(2) *T is non-separating on E_a; and*

(3) *there exists a monotone transformation (retracting) $W(A) = E_a$ which is non-separating and such that $T(x) = TW(x)$ on A.*

This theorem follows immediately from Theorem 6.1 and from Theorem 4.3 of J. F. Wardwell's paper.[13]

THEOREM 6.8. *If A is a dendrite and T is non-separating at a point,[14] then T is locally non-separating at p.*

If T were not locally non-separating at p, there would exist a neighborhood U_p such that for any neighborhood W_p lying in U_p there exist two points u and v of $W_p - T^{-1}(T(p))$ which do not lie in any connected set in $U_p - T^{-1}(T(p))$; that is, $T^{-1}(T(p))$ separates u and v in U_p. Since A is locally connected, it is possible to choose W_p such that it will contain only one component, namely, the one containing p. Thus there is an arc uv in W_p and this arc must contain points of $T^{-1}(T(p))$. Now since A is a dendrite, there is no other arc from u to v. Thus $T^{-1}(T(p))$ separates u and v in A, and this is contrary to T being non-separating at p. Therefore T is locally non-separating at p.

THEOREM 6.9. *If T is locally non-separating at all points of the dendrite A, then T is non-separating at all points of A.*

Let us first show that B is a point by assuming that it is not. If not, for all points x of B consider the closed sets $T^{-1}(x)$. Since A is a dendrite, the components of the closed sets $T^{-1}(x)$ are either points or continua which are dendrites. If any component of any set $T^{-1}(x)$ is a cut point of A, this fact will contradict the hypothesis that T is locally non-separating at that point. Thus no com-

[13] J. F. Wardwell, *Non-separating transformations*, loc. cit., p. 747.

[14] The transformation T will be called *non-separating at the point p* if $T^{-1}(T(p))$ does not separate A. It is to be noted that this is weaker than locally non-separating at p.

ponent of any set $T^{-1}(x)$ is a cut point. There remains the possibility that all cut points of A belong to components of the closed sets $T^{-1}(x)$ which are not points. If all cut points of A belong to one set $T^{-1}(x)$, then A belongs to $T^{-1}(x)$ since this set is closed and the only subset of A which is closed and contains all the cut points of A is A itself. But this is impossible under our assumption that B is not a point. Consider any two such components D_1 and D_2 belonging to different sets $T^{-1}(x)$. Let d_1 be a point of D_1 and d_2 be a point of D_2. Consider the arc d_1d_2. Since T is continuous and d_1d_2 is connected, $T(d_1d_2)$ is connected and contains more than one point. Thus $T(d_1d_2)$ contains uncountably many points, so d_1d_2 belong to uncountably many sets $T^{-1}(x)$. Hence uncountably many components of the sets $T^{-1}(x)$ are not points. But this is impossible, for if there were an uncountable number, there would be an uncountable number of diameter greater than some $\epsilon > 0$. But this is impossible since A is hereditarily locally connected.[15] Thus B is a point. Therefore $T^{-1}(B)$ does not separate A, and T is thus non-separating at all points.

Corollary. *If T is locally non-separating on the dendrite A, then B is a point.*

7. **Locally non-alternating transformations.**

A continuous transformation will be called *locally non-alternating at a point* p of $T^{-1}(x)$ if, for any neighborhood U_p of p, there exists a neighborhood V_p lying in U_p such that if u and v are any two points of $T^{-1}(y) \cdot V_p$, where y is any point of B distinct from x, then u and v lie in a connected subset of $U_p - T^{-1}(x)$. A continuous transformation will be called *locally non-alternating on a space* A, if it is locally non-alternating at all points of A.

Note that a non-alternating transformation is not necessarily locally non-alternating, and conversely. Also any homeomorphism is a locally non-alternating transformation. The space B under a locally non-alternating transformation does not necessarily have to be Peano and may contain cut points. Thus very few of the theorems of the last section (§6) carry over to the case of a locally non-alternating transformation.

Theorem 7.1. *In order that a transformation be locally non-alternating at a point p of $T^{-1}(x)$, it is necessary and sufficient that, for any neighborhood U_p of p, there exists a neighborhood W_p lying in U_p such that all the components of $W_p \cdot T^{-1}(y)$ belong to one component of $U_p - T^{-1}(x)$, for any $y \neq x$.*

Necessity. Let p be a point of some $T^{-1}(x)$ of A, and U_p any neighborhood of p. Since T is locally non-alternating, there exists a neighborhood W_p lying in U_p such that if u and v are any two points of the set $T^{-1}(y) \cdot W_p$ $(x \neq y)$, then u and v lie in a connected subset of $U_p - T^{-1}(x)$. Let u remain fixed and v range over all points of $T^{-1}(y) \cdot W_p$. Then each such point v lies with u in a connected subset C_v of $U_p - T^{-1}(x)$. Since each two sets C_v have the point u in common, $\sum C_v$ is connected and thus lies in a component of $U_p - T^{-1}(x)$.

[15] H. M. Gehman, *Concerning the subsets of a plane continuous curve*, Annals of Mathematics, vol. 27(1925), pp. 29–46, Theorem 5.

Sufficiency. Since all the components of $W_p \cdot T^{-1}(y)$ belong to one component of $U_p - T^{-1}(x)$, any two points of $W_p \cdot T^{-1}(y)$ will lie in a connected set in $U_p - T^{-1}(x)$.

THEOREM 7.2. *If T is locally non-alternating with A an arc and B a dendrite, and if $T^{-1}(x)$ contains no open set, then T is a homeomorphism and thus B is an arc.*

If T is a homeomorphism, the theorem is proved. If T is not a homeomorphism, there exist two points in A, a_0 and b_0, which go into the same point x_0 in B. Let α_0 be the arc a_0b_0. Not all the points of the arc α_0 go into x_0 in B; for if so, $T^{-1}(x_0)$ would contain an open set. Consider some point p_0 not of $T^{-1}(x_0)$ which lies in α_0. For any point b of B, $T^{-1}(b)$ is a closed set. Therefore on the arc α_0, consider any set $T^{-1}(x_1) \neq T^{-1}(x_0)$ which has points in α_0 from p_0 to a_0, and from p_0 to b_0. There would exist such a set; for if not, $T(a_0p_0)$ and $T(b_0p_0)$ would be Peano continua having only the two points $T(p_0)$ and x_0 in common. Then their sum would contain a simple closed curve. Call the first point of $T^{-1}(x_1)$ from p_0 to a_0, a_1; the first point of $T^{-1}(x_1)$ from p_0 to b_0, b_1. Call arc a_1b_1, α_1. By this procedure, α_1 lies entirely in α_0. By exactly the same method, we can pick an α_2 which lies entirely inside α_1. Continue this process indefinitely.

Let $\alpha_\omega = \prod_{i=1}^{\infty} \alpha_i$. If α_ω is a point, then T is not locally non-alternating at this point. If α_ω is an arc with end points a_ω and b_ω, then from the continuity of T it follows that $T(a_\omega) = T(b_\omega)$. Then we may form $\alpha_{\omega+1}$ just as above. Thus we continue this procedure and define the arcs α_β for all ordinals β of the first and second classes unless for some limiting number β, α_β is a point. But this is impossible; for at this point T would not be locally non-alternating since $a_i \to \alpha_\beta$ and $b_i \to \alpha_\beta$ and $T(a_i) = T(b_i)$.

But the process cannot continue through all ordinals of the first and second classes; for then A would contain an uncountable monotonic decreasing sequence of continua. This is an impossibility.

Then T is (1-1) and continuous from A to B. But as A and B are compact, T is then continuous from B to A. Thus T is a homeomorphism and B is an arc.

THEOREM 7.3. *If T is locally non-alternating where B is a dendrite and A is a dendrite in which the branch points are dense on no arc of A, and if $T^{-1}(x)$ contains no open set, then T is a homeomorphism.*

If T is a homeomorphism, the theorem is proved. If T is not a homeomorphism, there exist two points in A, a_0 and b_0, which go into the same point x_0 in B. Let α_0 be the arc a_0b_0. Not all the points of the arc α_0 go into x_0 in B; for if so, $T^{-1}(x_0)$ would contain α_0, and thus would contain an open set in α_0, since the branch points are not dense on α_0. The rest of the proof follows that of Theorem 7.2.

UNIVERSITY OF NEVADA.

REGULAR TRANSFORMATIONS

BY W. T. PUCKETT, JR.

1. Let M be a continuum and $T(M) = M'$ be an $(r - 1)$-regular transformation.[1] It is shown in this paper that if a', b' are any two points of M', then the s-dimensional Betti groups of $T^{-1}(a')$ and $T^{-1}(b')$ relative to M are isomorphic for $s = 0, 1, \cdots, r$. Furthermore, in case T is a monotone 0-regular transformation, it is shown that the 1-dimensional Betti group of M is the direct sum of two groups, one of which is isomorphic with the 1-dimensional Betti group of M', while the other is isomorphic with the 1-dimensional Betti group of $T^{-1}(a')$ relative to M for any point a' of M'. Thus $p^1(M) = p^1(M') + p^1(T^{-1}(a'), M')$, where $p^1(N)$ is the first Betti number of N and, for any point a' of M', $p^1(T^{-1}(a'), M)$ is the number of linearly independent cycles in $T^{-1}(a')$ relative to homologies in M.

The cycles and bounding relations used here are with respect to an arbitrary modulus $m \geqq 0$. The combinatorial notions used will be found in works of Alexandroff[2] and Vietoris.[3] After the convention of Alexandroff, $z^r \simeq 0 \pmod m$ indicates that the cycle z^r bounds an $(r + 1)$-dimensional complex relative to m, while $z^r \sim 0 \pmod m$ indicates that there exists a number α such that αz^r bounds. In case $m \geqq 2$, the two relations are the same.

2. Let the sequence of closed sets $\{A_n\}$, which are contained in a compact metric space M, converge to the limiting set A. The sequence is said to converge *r-regularly* (mod m) provided that for each $\epsilon > 0$ there exist positive numbers δ and N such that if $n > N$, any r-dimensional potentially bounding true cycle[4] (mod m) in A_n of diameter $< \delta$ is $\simeq 0 \pmod m$ in a subset of A_n of diameter $< \epsilon$. The convergence here defined differs from that given by G. T. Whyburn[5]

Received July 6, 1939.

[1] The transformation used here is a generalization of the 0-regular transformation defined by A. D. Wallace, Bulletin of the American Mathematical Society, abstract 44-3-161.

[2] *Dimensionstheorie*, Mathematische Annalen, vol. 106(1932), pp. 161–238. For elementary combinatorial notions see also P. Alexandroff and H. Hopf, *Topologie* I, Berlin, 1935.

[3] *Über den höheren Zusammenhang kompakter Räume* $\cdots$, Mathematische Annalen, vol. 97(1927), pp. 545–572.

[4] A true cycle $Z^r = (z_1^r, z_2^r, z_3^r, \cdots)$ is said to be potentially bounding provided all except a finite number of the z^r are potentially bounding (mod m). If r is > 0, any z^r is potentially bounding, while if $r = 0$, z^0 is potentially bounding if and only if the coefficient sum is $\equiv 0 \pmod m$.

[5] *On sequences and limiting sets*, Fundamenta Mathematicae, vol. 25(1935), pp. 408–426. It may be pointed out here that if the convergence is 0-regular for any m, it is 0-regular for every m, therefore is called simply 0-regular.

in that "complete" or "convergent" cycle has been replaced by "true" cycle. However, *for $m \geqq 2$, the two definitions are equivalent.* For let us assume that a sequence $\{A_n\}$ converges r-regularly (mod m) to A with respect to convergent cycles, but not with respect to true cycles. Then for some $\epsilon > 0$, any $\delta > 0$ and any N there exist an $n > N$ and a true cycle $Z_n^r = (z_1^r, \cdots, z_k^r, \cdots)$ of diameter $< \delta$ in A_n such that Z_n^r is not $\simeq 0$ in a subset of A_n of diameter $< \epsilon$. Hence there exists a positive number σ_n and a true cycle $Z_n'^r = (z_1'^r, \cdots, z_j'^r, \cdots)$ contained in Z_n^r such that each $z_j'^r$ is not $\underset{\sigma_n}{\simeq} 0$ in a subset of A_n of diameter $< \epsilon$. Since $m \geqq 2$, there exists a convergent[6] cycle $Z_n''^r = (z_1''^r, \cdots, z_i''^r, \cdots)$ contained in $Z_n'^r$, and by the above $Z_n''^r$ is not $\underset{\sigma_n}{\simeq} 0$ in a subset of A_n of diameter $< \epsilon$. Thus, since N may be taken arbitrarily large, $\{A_n\}$ does not converge r-regularly to A with respect to convergent cycles contrary to the hypothesis. Since any convergent cycle is a true cycle, the equivalence is established.

The continuous single-valued transformation $T(M) = M'$ is said to be *r-regular* (mod m) (see footnote 1) provided that for every sequence of points $\{a_n'\}$ converging to a point a' in M', the sequence $\{T^{-1}(a_n')\}$ converges s-regularly to $T^{-1}(a')$ for every $s \leqq r$. From the Eilenberg[7] characterization of an interior transformation[8] it follows that an r-regular transformation is interior. The following known characterization of an interior transformation is stated here for reference:

2.1. *Let M be compact and $T(M) = M'$ be continuous. In order that T be an interior transformation it is necessary and sufficient that for any $\epsilon > 0$ there exist a $\delta > 0$ such that for any point a' of M' and any point a of $T^{-1}(a')$, $T[U(a, \epsilon)]$ contains*[9] $U(a', \delta)$.[10]

The following is a characterization of an r-regular transformation:

2.2. *Let M be compact and $T(M) = M'$ be interior. In order that the transformation T be r-regular* (mod m) *it is necessary and sufficient that for a given $s \leqq r$ and $\epsilon > 0$ there exist a $\delta > 0$ such that for any point a' of M' and any s-dimensional potentially bounding true cycle Z^s of diameter $< \delta$ contained in $T^{-1}(a')$, $Z^s \simeq 0$* (mod m) *in a subset of $T^{-1}(a')$ of diameter $< \epsilon$.*

Proof. Suppose the conditions of the theorem are not necessary. Then for some $s \leqq r$ and some $\epsilon > 0$ there exist for each integer n a point a_n' of M' and an s-dimensional potentially bounding true cycle Z_n^s in $T^{-1}(a_n')$ of diameter $<$ $1/n$, which is not $\simeq 0$ in a subset of $T^{-1}(a_n')$ of diameter $< \epsilon$. It may be assumed $\{a_n'\}$ converges to a point a' of M' and $\{T^{-1}(a_n')\}$ converges to $T^{-1}(a')$. But as a consequence of the supposition this convergence is not s-regular,

[6] See *Dimensionstheorie*, loc. cit., p. 180.

[7] *Transformations en circonférence*, Fundamenta Mathematicae, vol. 24(1935), p. 174.

[8] A continuous transformation is interior provided any open set in the original space transforms into an open set in the image space.

[9] By $U(a, \epsilon)$ is meant the ϵ-neighborhood of a.

[10] See Whyburn, *The mapping of Betti groups under interior transformations*, this Journal, vol. 4(1938), p. 1.

contrary to the hypothesis that T is an r-regular transformation. Suppose the conditions of the theorem are not sufficient. Then there exists a sequence of points $\{a'_n\}$ converging to a point a' of M' such that $T^{-1}(a'_n)$ does not converge s-regularly to $T^{-1}(a')$ for some $s \leqq r$. Thus it may be assumed that for each n there exists an s-dimensional potentially bounding true cycle Z^s_n of diameter $< 1/n$ in $T^{-1}(a'_n)$ which is not $\simeq 0$ in a subset of $T^{-1}(a'_n)$ of diameter $< \epsilon$. This contradicts the existence of the δ given by the hypotheses, and the theorem is established.

2.3. *Let M be compact, $T(M) = M'$ be an $(r - 1)$-regular* (mod m) *transformation, and a' be any point of M'. If $\epsilon > 0$ there exists a $\delta > 0$ such that to each r-dimensional δ-cycle z^r in $T^{-1}(a')$ there corresponds an r-dimensional true cycle Z^r with the property that $z^r \underset{\epsilon}{\simeq} Z^r$* (mod m) *in $T^{-1}(a')$.*

Proof. Let $d_r < \epsilon/(r + 1)$ and let $2d_{r-1}$ be a number given by 2.2 such that every $(r - 1)$-dimensional potentially bounding true cycle in $T^{-1}(a')$, for any point a' of M', of diameter $< 2d_{r-1}$ is $\simeq 0$ in a subset of $T^{-1}(a')$ of diameter $< d_r$. Generally, let $2d_k$ be a number given by 2.2 such that every k-dimensional potentially bounding true cycle in $T^{-1}(a')$ of diameter $< 2d_k$ is $\simeq 0$ in a subset of $T^{-1}(a')$ of diameter $< d_{k+1}$. Finally, let d_0 be a number given by 2.2 such that every 0-dimensional potentially bounding true cycle in $T^{-1}(a')$ of diameter $< d_0$ is $\simeq 0$ in a subset of $T^{-1}(a')$ of diameter $< d_1$. For $\delta = d_0$ let z^r be any r-dimensional δ-cycle in $T^{-1}(a')$, then it will be shown that there exists a true cycle $Z^r \underset{\epsilon}{\simeq} z^r$ (mod m) in $T^{-1}(a')$.

Assume $z^r = \sum_{i=0}^{\mu_r} \alpha^i \Delta^r_i$, where the Δ^r_i are r-dimensional simplexes of z^r. Generally, for any $s \leqq r$ let Δ^s_i $(i = 0, 1, \cdots, \mu_s)$ be the s-dimensional simplexes in z^r and let a^s_i be a point of $T^{-1}(a') \cdot U(|\Delta^s_i|, \delta)$, which may be a vertex of z^r. For each Δ^1_i there exists a null sequence of complexes $\{{}_nK^1_i\}$ contained in a subset of $T^{-1}(a')$ of diameter $< d_1$, such that[11] ${}_n\dot{K}^1_i = \dot{\Delta}^1_i$. Define the cycles

$$ {}_nx^1_i = {}_nK^1_i - \Delta^1_i \qquad (i = 0, 1, \cdots, \mu_1;\ n = 1, 2, \cdots) $$

and the $(d_0 + d_1)$-complexes ${}_nQ^2_i = (a^1_i\ {}_nx^1_i)$.[12] Now with each $\dot{\Delta}^2_i = \sum_{j=0}^{2} \beta^{ij} \Delta^1_{i_j}$ there is associated a true cycle $\{{}_ny^1_i\} = \{\sum_{j=0}^{2} \beta^{ij}\ {}_nK^1_{i_j}\}$, which is of diameter $< 2d_1$. Hence there exists a null sequence of complexes $\{{}_nK^2_i\}$ contained in a subset of $T^{-1}(a')$ of diameter $< d_2$ such that, for each n, ${}_n\dot{K}^2_i = {}_ny^1_i$. Define the $(d_0 + d_1)$-complexes ${}_nL^2_i = \sum_{j=0}^{2} \beta^{ij}\ {}_nQ^2_{i_j}$, the cycles ${}_nx^2_i = {}_nK^2_i - \Delta^2_i - {}_nL^2_i$,

[11] If K is a complex and m is given, then $\dot{K}$ denotes the boundary of K (mod m).

[12] If $\Delta^r = (a_0, a_1, \cdots, a_r)$ is an r-dimensional oriented simplex and a is a point, then $(a\Delta^r)$ is the $(r + 1)$-dimensional oriented simplex $(a, a_0, a_1, \cdots, a_r)$; and if $K^r = \sum \alpha^i \Delta^r_i$, then $(aK^r) = \sum \alpha^i (a\Delta^r_i)$. It is known that $(aK^r)^{\cdot} = K^r + (a\dot{K}^r)$.

and the $(d_0 + d_1 + d_2)$-complexes ${}_nQ_i^3 = (a_i^2\, {}_nx_i^2)$. Generally, for $k \leqq r$ there is associated with each $\dot{\Delta}_i^k = \sum_{j=0}^{k} \beta^{i_j}\Delta_{i_j}^{k-1}$ $(i = 0, 1, \cdots, \mu_k)$ a true cycle $\{{}_ny_i^{k-1}\} = \{\sum_{j=0}^{k} \beta^{i_j}\, {}_nK_{i_j}^{k-1}\}$ which is of diameter $< 2d_{k-1}$. Hence there exists a null sequence of complexes $\{{}_nK_i^k\}$ contained in a subset of $T^{-1}(a')$ of diameter $< d_k$ such that, for each n, ${}_n\dot{K}_i^k = {}_ny_i^{k-1}$. Define the $(\sum_{j=0}^{k-1} d_j)$-complexes ${}_nL_i^k = \sum_{j=0}^{k} \beta^{i_j}\, {}_nQ_{i_j}^k$, the cycles ${}_nx_i^k = {}_nK_i^k - \Delta_i^k - {}_nL_i^k$, and the $(\sum_{j=0}^{k} d_j)$-complexes ${}_nQ_i^{k+1} = (a_i^k\, {}_nx_i^k)$. Thus, for each n, ${}_nQ^{r+1} = \sum_{i=0}^{\mu_r} \alpha^i\, {}_nQ_i^{r+1}$ is an ϵ-complex, since $\sum_{i=0}^{r} d_i < (r + 1)d_r \leqq \epsilon$. Moreover, it will be shown that, for each n, ${}_n\dot{Q}^{r+1} = \sum_{i=0}^{\mu_r} \alpha^i\, {}_nK_i^r - z^r$. Now

$$
\begin{aligned}
{}_n\dot{Q}^{r+1} &= \sum_{i=0}^{\mu_r} \alpha^i({}_nx_i^r) = \sum_{i=0}^{\mu_r} \alpha^i\, {}_nK_i^r - \sum_{i=0}^{\mu_r} \alpha^i \Delta_i^r - \sum_{i=0}^{\mu_r} \alpha^i\, {}_nL_i^r \\
&= \sum_{i=0}^{\mu_r} \alpha^i\, {}_nK_i^r - z^r - \sum_{j=0}^{\mu_{r-1}} \lambda^j\, {}_nQ_j^r,
\end{aligned}
$$

where λ^j is the collected coefficient of ${}_nQ_j^r$ as it occurs in the various ${}_nL_i^r$. But for each n the complex ${}_nQ_j^r$ occurs in each $\alpha^i\, {}_nL_i^r$ with the same coefficient as Δ_j^{r-1} occurs in $(\alpha^i \Delta_i^r)^{\cdot}$. Hence, since each Δ_j^{r-1} occurs in $\dot{z}^r = \sum_{i=0}^{\mu_r} (\alpha^i\Delta_i^r)^{\cdot}$ with coefficient 0, each $\lambda^j = 0$. Consequently ${}_nz^r = \sum_{i=0}^{\mu_r} \alpha^i\, {}_nK_i^r$ is a cycle and ${}_n\dot{Q}^{r+1} = {}_nz^r - z^r$. Therefore, since $\{{}_nK_i^r\}$ is a null sequence for each $i = 0, 1, \cdots, \mu_r$, $Z^r = \{{}_nz^r\}$ is the true cycle sought.

On account of the convergence theorem (see footnote 6) one has

2.31. *For* $m \geqq 2$, 2.3 *yields a convergent cycle* $Z^r \simeq z^r$ (mod m) *in* $T^{-1}(a')$.

3. The *r-dimensional homology*[13] or *Betti group* (mod m) *of* M is the group $B_m^r(M)$, whose elements are classes of homologous convergent cycles; i.e., if ζ is an element of $B_m^r(M)$ and Z_1^r, Z_2^r are two convergent cycles in ζ, then there exist α^i not $\equiv 0$ (mod m) such that $\alpha^1 Z_1^r + \alpha^2 Z_2^r \sim 0$ (mod m) in M. The least number, $p_m^r(M)$, of elements which form a basis[14] for $B_m^r(M)$ is called the *r-th Betti number* (mod m) *of* M. As indicated by Vietoris[14] this number is at most countably infinite.

Let N be a closed subset of M. The *r-dimensional Betti group* (mod m) *of* N *relative to* M, $B_m^r(N, M)$, is composed of the classes of r-dimensional convergent

[13] See Vietoris, loc. cit., p. 459.

[14] See Vietoris, loc. cit., pp. 463–464.

cycles in N relative to homologies in M; i.e., if ζ is an element of $B^r_m(N, M)$ and Z^r_1, Z^r_2 are cycles of ζ, then there exist α^i not $\equiv 0 \pmod m$ such that $\alpha^1 Z^r_1 + \alpha^2 Z^r_2 \sim 0 \pmod m$ in M. The least number of elements, $p^r_m(N, M)$, which form a basis in $B^r_m(N, M)$ is called the *r-th Betti number of N relative to M*. Clearly $p^r_m(N, M) \leqq p^r_m(N)$ and $p^r_m(N, M) \leqq p^r_m(M)$.

3.1. THEOREM. *Let $T(M) = M'$ be an $(r - 1)$-regular* (mod m) *transformation, where M is a continuum.*[15] *If a', b' are any two points of M', then $B^s_m(T^{-1}(a'), M)$ is isomorphic with $B^s_m(T^{-1}(b'), M)$ for $s \leqq r$.*

Since T is $(s - 1)$-regular for every $s \leqq r$, it suffices to prove the theorem for $s = r$. This is accomplished by the following sequence of assertions:

3.2. *For any $\epsilon > 0$ there exists a $\delta > 0$ such that to each r-dimensional δ-cycle z^r in $T^{-1}(a')$ there corresponds a δ-cycle y^r in $T^{-1}(b')$ such that $y^r \underset{\epsilon}{\simeq} z^r$* (mod m) *in M.*

Proof. For the given $\epsilon > 0$ let $3\delta \leqq \epsilon$ be a number satisfying 2.3. For the number $\delta > 0$, let σ be a number given by 2.1, such that, for each point c of M, $T[U(c, \delta)]$ contains $U[T(c), \sigma]$. Let $a' = a'_0, a'_1, \cdots, a'_\mu = b'$ be a σ-chain joining a' and b' in M'. It will be shown that if y^r_i is a δ-cycle in $T^{-1}(a'_i) = A_i$, then there exists a δ-cycle y^r_{i+1} $(i = 0, 1, \cdots, \mu - 1)$ in $T^{-1}(a'_{i+1}) = A_{i+1}$ such that $y^r_{i+1} \underset{\epsilon}{\simeq} y^r_i$, and the assertion follows.

Let $(b_1, \cdots, b_j, \cdots, b_\nu)$ be the vertices of y^r_i. Then to each b_j there corresponds a point c_j of A_{i+1} such that $\rho(b_j, c_j) < \delta$, because of the choice of σ. Suppose $y^r_i = \sum \alpha^k \Delta^r_k$, where $\Delta^r_k = (b_{j_0}, \cdots, b_{j_r})$ are the simplexes of y^r_i. Then $z^r = \sum \alpha^k D^r_k$, where $D^r_k = (c_{j_0}, \cdots, c_{j_r})$, is a 3δ-cycle and $z^r \underset{3\delta}{\simeq} y^r_i$.[16] Thus, from 2.3 and the choice of 3δ it follows that there exists a true cycle $Z^r = \{z^r_n\}$ in A_{i+1} such that $Z^r \underset{\epsilon}{\simeq} z^r$. Consequently, for sufficiently large n, z^r_n is a δ-cycle and is $\underset{\epsilon}{\simeq} z^r$. Define $y^r_{i+1} = z^r_n$; then $y^r_{i+1} \underset{\epsilon}{\simeq} z^r \underset{3\delta}{\simeq} y^r_i$, i.e., $y^r_{i+1} \underset{\epsilon}{\simeq} y^r_i$, since $3\delta \leqq \epsilon$.

3.3. *If $Z^r = (z^r_1, \cdots, z^r_n, \cdots)$ is an r-dimensional true cycle in $T^{-1}(a')$, then there exists an r-dimensional true cycle $Y^r = (y^r_1, \cdots, y^r_n, \cdots)$ in $T^{-1}(b')$ such that $Y^r \simeq Z^r$* (mod m) *in M, and if Z^r is a convergent cycle so also is Y^r.*

Proof. Let $\{\epsilon_i\}$ be a sequence of positive numbers converging to zero. For each ϵ_i, let δ_i be a number given by 3.2. It may be assumed each z^r_n of Z^r is a δ_1-cycle. Moreover, for each i there exists an integer N_i such that, for $n > N_i$, z^r_n is a δ_i-cycle. From (3.2) it follows that for $n \leqq N_2$ there exists a δ_1-cycle y^r_n in $T^{-1}(b')$ which is $\underset{\epsilon_1}{\simeq} z^r_n$ in M. Generally, for each i and n such that $N_i < n \leqq N_{i+1}$ there exists a δ_i-cycle y^r_n in $T^{-1}(b')$ which is $\underset{\epsilon_i}{\simeq} z^r_n$ in M. Then $Y^r = (y^r_1, \cdots, y^r_n, \cdots)$ is the true cycle sought.

3.4. *Let $Z^r_1, \cdots, Z_q$, $q = p^r_m(T^{-1}(a'), M)$, be a basis in $B^r_m(T^{-1}(a'), M)$. If, for each i, Y^r_i is a convergent cycle in $T^{-1}(b')$ corresponding to Z^r_i given by* (3.3), *then the Y^r_i form a basis in $B^r_m(T^{-1}(b'), M)$.*

[15] I.e., a closed compact connected set.

[16] This is essentially a part of a proof given by Vietoris, loc. cit., (3).

Proof. In the first place, the Y_i^r are linearly independent relative to homologies in M. For suppose there exist α^i not all $\equiv 0 \pmod m$ such that $\sum_{i=1}^{q} \alpha^i Y_i^r \sim 0$. Then $\sum_{i=1}^{q} \alpha^i Z_i^r \sim 0$ in M, contrary to the hypothesis that the Z_i^r form a basis in $B_m^r(T^{-1}(a'), M)$. Now let Y_0^r be any r-dimensional convergent cycle in $T^{-1}(b')$ and let Z_0^r be a convergent cycle in $T^{-1}(a')$ given by 3.3 such that $Y_0^r \simeq Z_0^r$ in M. Then there exist β^i not all $\equiv 0 \pmod m$ such that $\sum_{i=0}^{q} \beta^i Z_i^r \sim 0$ in M. Hence $\sum_{i=0}^{q} \beta^i Y_i^r \sim 0$ in M, and consequently the convergent cycles $Y_1^r, \cdots, Y_q^r$ form a basis in $B_m^r(T^{-1}(b'), M)$.

3.5. Corollary. *Let $T(M) = M'$ be an $(r-1)$-regular* (mod m) *transformation, where M is a continuum. If a', b' are any two points of M', then $p_m^r(T^{-1}(a'), M) = p_m^r(T^{-1}(b'), M)$.*

4. Theorem. *Let M be a continuum and $T(M) = M'$ be a monotone*[17] *0-regular*[18] *transformation. If Z^1 is a 1-dimensional true cycle* (mod m) *in M such that $T(Z^1) = Z' \simeq 0$ in M', then for each point a' in M' there exists a true cycle Y^1 in $T^{-1}(a')$ such that $Y^1 \simeq Z^1$ in M.*

Let $\epsilon > 0$ be arbitrary but fixed, and let a' be any point of M'. By 2.2 there exists a number e such that any 0-dimensional potentially bounding true cycle of diameter $< 3e$ contained in $T^{-1}(a')$ is $\simeq 0$ in a subset of $T^{-1}(a')$ of diameter $< \frac{1}{2}\epsilon$. Moreover, by 2.1, there exists a $\delta > 0$ such that $T[U(a, e)]$ contains $U[T(a), \delta]$ for every point a of M. Finally, let z^1 be a 1-dimensional e-cycle of M such that if a, b are vertices of a simplex of z^1, then $\rho[T(a), T(b)] < \delta$. Under these conditions the following assertions are established.

4.1. *If the two simplexes $\Delta_0 = (a_0, b_0)$ and $\Delta = (a, b)$ of z^1 are such that $T(\Delta_0) = T(\Delta)$ while $T(a_0)$ is not $T(b_0)$, then there exists a 1-dimensional e-cycle x^1 in M such that* (i) $x^1 \underset{\epsilon}{\simeq} z^1$, (ii) *$T(x^1) = x'$ differs from $T(z^1) = z'$ only by a degenerate cycle, and* (iii) *x^1 does not contain Δ.*

Proof. It may be assumed that $T(a_0) = T(a) = a'$. Since T is monotone, there exists an e-chain $a_0, a_1, \cdots, a_\mu, a_{\mu+1} = a$ in $T^{-1}(a')$. Let $T(b_0) = T(b) = b'$, then $\rho(a', b') < \delta$, by the conditions on z^1. Hence on account of the choice of δ, there exists for each a_j a point b_j of $T^{-1}(b')$ such that $\rho(a_j, b_j) < e$. Consequently, $b_0, b_1, \cdots, b_\mu, b_{\mu+1} = b$ is a $3e$-chain in $T^{-1}(b')$. Let $z^1 = \alpha\Delta + \sum_{i=0}^{\lambda} \alpha^i \Delta_i$ and define

$$C_j = \alpha l_j = \alpha(a_j, a_{j+1}) \qquad (j = 0, 1, \cdots, \mu),$$

[17] A continuous transformation is monotone if the inverse set of each point of the image space is connected.

[18] A transformation which is 0-regular for any m is 0-regular for every m, therefore is called simply a 0-regular transformation.

and $s_j = (b_j, b_{j+1})$. Now $(\alpha s_j)^{\cdot}$ is a potentially bounding 0-dimensional true cycle of diameter $< 3e$. Thus it is $\simeq 0$ in a subset of $T^{-1}(b')$ of diameter $< \frac{1}{2}\epsilon$, because of the choice of e. Therefore, there exists, for each $j = 0, 1, \cdots, \mu$, an e-complex Q_j of diameter $< \frac{1}{2}\epsilon$ such that $\dot{Q}_j = (\alpha s_j)^{\cdot}$. Define $K^2 = \sum_{j=0}^{\mu} [(a_j Q_j) - (b_{j+1} C_j)]$, then K^2 is an ϵ-complex and

$$\begin{aligned}\dot{K}^2 &= \sum_{j=0}^{\mu} [Q_j - C_j + (a_j \dot{Q}_j) - (b_{j+1} \dot{C}_j)] \\ &= \sum_{j=0}^{\mu} [Q_j - C_j + \{a_j(\alpha b_{j+1} - \alpha b_j)\} - \{b_{j+1}(\alpha a_{j+1} - \alpha a_j)\}] \\ &= \sum_{j=0}^{\mu} [Q_j - C_j] - \alpha\Delta_0 + \alpha\Delta.\end{aligned}$$

Consequently $x^1 = z^1 - \dot{K}^2$ is the desired cycle.

4.2. *Let z^1 have the property that $T(z^1) = z' = \dot{K}'$, where K' is a 2-dimensional δ-complex. If $\Delta' = (T(a), T(b), c')$ is a non-degenerate simplex of K' such that a and b are vertices of a simplex of z^1, then there exists an e-cycle x^* with the properties* (i) $x^* \underset{\epsilon}{\simeq} z^1$ *in M and* (ii) $T(x^*) = x^{*\prime}$ *differs from $(K' - \alpha\Delta')^{\cdot}$, where α is the coefficient of Δ' in K', only by a degenerate cycle.*

Proof. Since $\rho(T(a), c') < \delta$ and $\rho(T(b), c') < \delta$, there exist points c_a and c_b of $T^{-1}(c')$ such that $\rho(c_a, a) < e$ and $\rho(c_b, b) < e$, and consequently $\rho(c_a, c_b) < 3e$. Thus, just as in the proof of 4.1, an e-complex Q of diameter $< \frac{1}{2}\epsilon$ can be found in $T^{-1}(c')$ such that $\dot{Q} = (\alpha s)^{\cdot}$, where $s = (c_a, c_b)$, because of the choice of e. Let us define $L = (aQ) + \alpha(a, b, c_b)$ and $x^* = z^1 - \dot{L}$. Then x^* is the desired cycle; for L is an ϵ-complex and

$$\begin{aligned}T(x^*) = T(z^1 - \dot{L}) = T(z^1) - T(\dot{L}) &= T(z^1) - [T(L)]^{\cdot} \\ &= \dot{K}' - [\alpha\Delta' - \text{degenerate complex}]^{\cdot} \\ &= (K' - \alpha\Delta')^{\cdot} + (\text{degenerate complex})^{\cdot} \\ &= (K' - \alpha\Delta')^{\cdot} + \text{degenerate cycle}.\end{aligned}$$

4.3. *If z^1 is such that $T(z^1) = z' \underset{\delta}{\simeq} 0$ and a' is any point of M', then there exists an e-cycle y^1 in $T^{-1}(a')$ such that $y^1 \underset{\epsilon}{\simeq} z^1$ in M.*

Proof. By successive applications of 4.1 an e-cycle x_1 is obtained such that (i) $x_1 \underset{\epsilon}{\simeq} z^1$ in M, (ii) $T(x_1) = T(z^1) +$ a degenerate cycle, and (iii) if Δ_1 and Δ_2 are two fundamental simplexes of x_1, then $T(\Delta_1) = T(\Delta_2)$ if and only if both are degenerate, and therefore map into the same point. Since $z' \underset{\delta}{\simeq} 0$, there exists a δ-complex K' in M' such that $T(x_1) = \dot{K}'$. Let

$$K' = \sum_{i=1}^{\mu} \alpha^i \Delta_i + \text{degenerate complex},$$

where each Δ_i is a non-degenerate 2-dimensional simplex. Moreover, it may be assumed Δ_1 has the form $\Delta_1 = (T(a), T(b), c')$, where a and b are vertices of a simplex of x_1. By 4.2 there exists an e-cycle x_1^* such that (i) $x_1^* \underset{\epsilon}{\simeq} x_1$ in M and (ii) $T(x_1^*) = (K' - \alpha^1\Delta_1)^{\cdot} +$ degenerate cycle. Continue the application of 4.1 and 4.2 to obtain finally an e-cycle x_μ^* such that (i) $x_\mu^* \underset{\epsilon}{\simeq} x_\mu \underset{\epsilon}{\simeq} x_{\mu-1}^* \underset{\epsilon}{\simeq} x_{\mu-1} \underset{\epsilon}{\simeq} \cdots \underset{\epsilon}{\simeq} x_1^* \underset{\epsilon}{\simeq} x_1 \underset{\epsilon}{\simeq} z^1$ and (ii) $T(x_\mu^*)$ is a degenerate cycle, since it is the boundary of a degenerate complex. Hence each simplex in $T(x_\mu^*)$ must be a reiteration of some one of a finite set of points $a_1', a_2', \cdots, a_\lambda'$. Thus x_μ^* has the form $x_\mu^* = z_1 + z_2 + \cdots + z_\lambda$, where z_j is contained in $T^{-1}(a_j')$. But by 3.2 it may be assumed that ϵ and e are so related that for each j there exists an e-cycle y_j^1 in $T^{-1}(a')$ such that $y_j^1 \underset{\epsilon}{\simeq} z_j$ in M. Therefore, $y^1 = y_1^1 + y_2^1 + \cdots + y_\lambda^1$ is the cycle sought, for $y^1 \underset{\epsilon}{\simeq} x_\mu^* \underset{\epsilon}{\simeq} z^1$.

Proof of theorem. Let $\{\epsilon_i\}$ be a null sequence of positive numbers and let $\{e_i\}$ and $\{\delta_i\}$ be related to the corresponding ϵ_i as above. In $Z^1 = \{z_n^1\}$ it may be assumed that, for all n, z_n^1 is a cycle satisfying 4.3 for $(\epsilon_1, e_1, \delta_1)$. Generally, there exists an integer N_i such that, for $n > N_i$, z_n^1 is a cycle satisfying 4.3 for $(\epsilon_i, e_i, \delta_i)$. Thus 4.3 gives a true cycle $Y^1 = \{y_n^1\}$ such that, for $N_i < n \leqq N_{i+1}$, y_n^1 is an e_i-cycle and $y_n^1 \underset{\epsilon_i}{\simeq} z_n^1$ in M. Therefore, since both $\{\epsilon_i\}$ and $\{e_i\}$ converge to zero, it follows that Y^1 is the true cycle sought.

In the same way one can establish the following:

4.4. Theorem. *If under the conditions of 4 $T(Z^1) \sim 0$ in M, then there exists a true cycle Y^1 in $T^{-1}(a')$ such that $Y^1 \sim Z^1$ in M.*

5. Theorem. *If M is a continuum and $T(M) = M'$ is a monotone 0-regular transformation, then $B_m^1(M) = U + V$, where U is isomorphic with $B_m^1(M')$ and V is isomorphic with $B_m^1(T^{-1}(a'), M)$ for any point a' of M'.*

Proof. Let $Z_1', \cdots, Z_{p'}'$, $p' = p_m^1(M')$, be a basis for $B_m^1(M')$; then according to a theorem of Vietoris[19] there exist convergent cycles $Z_1, \cdots, Z_{p'}$ of M such that $T(Z_i) \sim Z_i'$. Let $Z_1^*, \cdots, Z_q^*$, $q = p_m^1(T^{-1}(a'), M)$, be a basis for $B_m^1(T^{-1}(a'), M)$. The convergent cycles $Z_1, \cdots, Z_{p'}, Z_1^*, \cdots, Z_q^*$ form a linearly independent (mod m) set of cycles. For suppose

$$\sum_{i=1}^{p'} \alpha^i Z_i + \sum_{j=1}^{q} \beta^j Z_j^* \sim 0$$

in M. Then

$$T\left(\sum_{i=1}^{p} \alpha^i Z_i + \sum_{j=1}^{q} \beta^j Z_j^*\right) \sim \sum_{i=1}^{p'} \alpha^i Z_i' \sim 0.$$

Thus $\alpha^i \equiv 0 \pmod m$, since the Z_i' form a basis in $B_m^1(M')$. Hence $\sum_{j=1}^{q} \beta^j Z_j^* \sim 0$ in M, and consequently $\beta^j \equiv 0 \pmod m$, since the Z_j^* form a basis in

[19] See Vietoris, loc. cit., pp. 468–469. While Vietoris uses $m = 0$ or 2 only, his proof is valid for any $m \geqq 0$.

$B^1_m(T^{-1}(a'), M)$. Now let Z^1 be any 1-dimensional convergent cycle in M, then there exist α^i such that $\alpha^0 T(Z^1) + \sum_{i=1}^{p'} \alpha^i Z'_i \sim 0$, since the Z'_i form a basis in $B^1_m(M')$. Now $X^1 = \alpha^0 Z^1 + \sum_{i=1}^{p'} \alpha^i Z_i$ is a cycle in M such that $T(X^1)$ is ~ 0 in M'. Therefore by 4.4 there exists a cycle Y^1 in $T^{-1}(a')$ such that $Y^1 \sim X^1$ in M. But there exist β^j such that $\beta^0 Y^1 + \sum_{j=1}^{q} \beta^j Z^*_j \sim 0$, since the Z^*_j form a basis in $B^1_m(T^{-1}(a'), M)$. Hence

$$0 \sim \beta^0 Y^1 + \sum_{j=1}^{q} \beta^j Z^*_j \sim \beta^0 X^1 + \sum_{j=1}^{q} \beta^j Z^*_j = \beta^0 \alpha^0 Z^1 - \beta^0 \sum_{i=1}^{p'} \alpha^i Z_i + \sum_{j=1}^{q} \beta^j Z^*_j ,$$

i.e., there exist λ's and β's such that

$$\lambda^0 Z^1 + \sum_{i=1}^{p'} \lambda^i Z_i + \sum_{j=1}^{q} \beta^j Z^*_j \sim 0.$$

5.1. Corollary. *If M is a continuum and $T(M) = M'$ is a monotone 0-regular transformation, then $p^1_m(M) = p^1_m(M') + p^1_m(T^{-1}(a'), M)$, where a' is any point of M'.*

University of Virginia.

SOME GENERALIZATIONS OF THE THEORY OF ORTHOGONAL POLYNOMIALS

By Glenn Peebles

1. **Introduction.** If a function $\rho(x)$ is integrable and non-negative on an interval (a, b) and is such that $\int_a^b \rho(x)\,dx > 0$, a set of polynomials $[p_n(x) = a_n x^n + b_n x^{n-1} + \cdots]$ is uniquely determined, except for a constant factor, by the relations

$$\int_a^b \rho(x) p_n(x) p_m(x)\,dx = 0, \qquad m \neq n. \tag{1}$$

The set of polynomials defined in this way is said to be orthogonal with respect to the weight function $\rho(x)$ over the interval (a, b).

For the weight functions

$$\begin{aligned} &\rho(x) = (x-a)^\alpha (b-x)^\beta, && \alpha > -1, \beta > -1,\\ &\rho(x) = (x-a)^\alpha e^{-\beta x}, && \alpha > -1, \beta > 0, b = \infty,\\ &\rho(x) = e^{-\alpha x^2 + \beta x}, && \alpha > 0, a = -\infty, b = \infty, \end{aligned} \tag{a}$$

the polynomials $[p_n(x)]$ are respectively those of Jacobi, Laguerre, and Hermite. Each of these weight functions satisfies the Pearson differential equation

$$\frac{1}{\rho(x)} \frac{d}{dx} \rho(x) = \frac{Ax + B}{Cx^2 + Dx + E} \equiv \frac{Ax + B}{M(x)}, \tag{2}$$

if A, B, C, D, E are given suitable values.

The class of functions defined by (2) when A, B, C, D, E range through all real values $(Cx^2 + Dx + E \not\equiv 0)$ is such that for each non-identically vanishing member $\rho(x)$, the expression $[\rho(x)]^{-1} d^n[M^n(x)\rho(x)]/dx^n$; where $M^n(x)$ means $[M(x)]^n$, is a polynomial in x of degree n at most. The set of polynomials

$$q_0(x) = 1, \qquad q_n(x) = \frac{1}{\rho(x)} \frac{d^n}{dx^n} [M^n(x)\rho(x)] \qquad (n = 1, 2, 3, \cdots) \tag{3}$$

satisfies (1) when $\rho(x)$ is one of the functions (a). In other cases the property of orthogonality is lost because $\rho(x)$ is such that the integral (1) ceases to have a meaning. The corresponding sets of polynomials (3) will by way of distinction be called non-orthogonal.

As is known, each of the systems of orthogonal polynomials satisfies a recursion formula and a Christoffel-Darboux identity deduced from the recursion formula, and has the property of representing suitable functions by means of convergent

Received July 17, 1939.

series. These considerations suggest the question: To what extent do these characteristics persist in the case of the non-orthogonal systems defined by (3)?[1] To answer this question (at least in part) is the purpose of the first part of this paper.

The second part is concerned with a generalization of the ordinary concept of orthogonal polynomials, namely, the construction of such polynomials *for a class of weight functions which change sign in the interval of orthogonality.*[2] A theorem is given which, under the assumption of the existence of the polynomials $[q_n(x)]$ orthogonal with respect to a weight function $\rho(x)$, places on a general polynomial $\pi_m(x)$ a necessary and sufficient condition for the existence of the polynomials orthogonal with respect to the weight function $\pi_m(x)\rho(x)$. The necessary and sufficient condition is formulated in terms of a set of general determinants which include, as special cases, Vandermonde and Wronskian determinants of the polynomials $[q_n(x)]$.

2. **The recursion formula and the Christoffel-Darboux identity.** If the factor in $p_n(x)$ left undetermined by (1) is so taken that $\int_a^b \rho(x)[p_n(x)]^2\,dx = 1$, the p's satisfy the recursion formula

$$xp_n(x) = \frac{a_n}{a_{n+1}}\,p_{n+1}(x) + \left[\frac{b_n}{a_n} - \frac{b_{n+1}}{a_{n+1}}\right]p_n(x) + \frac{a_{n-1}}{a_n}\,p_{n-1}(x). \tag{4}$$

For the moment let A, B, C, D, and E in (2) be confined to those ranges for which $\rho(x)$ takes the form $(x-a)^\alpha(b-x)^\beta$, $\alpha > -1$, $\beta > -1$, and let (4) be written in terms of the q's by placing $p_n(x) = q_n(x)/I_n$, $a_n = \alpha_n/I_n$, $b_n = \beta_n/I_n$, where $q_n(x)$ is defined by (3), α_n and β_n are the first and second leading coefficients of $q_n(x)$, and

$$I_n = \left\{\int_a^b (x-a)^\alpha(b-x)^\beta[q_n(x)]^2\,dx\right\}^{\frac{1}{2}}.$$

Then (4) may be reduced to the form

$$xq_n(x) = \frac{\alpha_n}{\alpha_{n+1}}\,q_{n+1}(x) + \left[\frac{\beta_n}{\alpha_n} - \frac{\beta_{n+1}}{\alpha_{n+1}}\right]q_n(x) + \frac{\alpha_{n-1}}{\alpha_n}\,\frac{I_n^2}{I_{n-1}^2}\,q_{n-1}(x). \tag{5}$$

[1] A number of the recurrence formulas for the Jacobi, Laguerre, and Hermite polynomials have been shown to exist for the non-orthogonal polynomials (3) by E. H. Hildebrandt, *Systems of polynomials connected with the Charlier expansions and the Pearson differential and difference equations*, Annals of Mathematical Statistics, vol. 2(1931), pp. 379–439; pp. 393–411. For other properties of the polynomials (3), see also F. S. Beale, *On the polynomials related to the differential equation* $\frac{1}{y}\frac{dy}{dx} = \frac{a_0 + a_1 x}{b_0 + b_1 x + b_2 x^2} \equiv \frac{N}{D}$, Annals of Mathematical Statistics, vol. 8(1937), pp. 206–233.

[2] On this topic, see Jacques Chokhatte (J. Shohat), *Sur les fractions continues algébriques*, Comptes Rendus, vol. 191(1930), pp. 474–475; J. Shohat, *Sur les polynomes orthogonaux généralisés*, Comptes Rendus, vol. 207(1938), pp. 556–558; A. Tartler, *On a certain class of orthogonal polynomials*, Amer. Jour. Math., vol. 57(1935), pp. 627–644; G. Szegö, *Über gewisse orthogonale Polynome, die zu einer oszillierenden Belegungsfunktion gehören*, Mathematische Annalen, vol. 110 (1935), pp. 501–513.

Simple but lengthy calculations show that

$$
\begin{aligned}
&\alpha_0 = 1,\\
&\alpha_n = [A + (n+1)C][A + (n+2)C]\cdots[A + 2nC] \qquad (n > 0),\\
(6)\qquad &\beta_n = n[A + (n+1)C][A + (n+2)C]\cdots[A + (2n-1)C][B + nD]\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (n > 0),\\
&\frac{I_n^2}{I_{n-1}^2} = \frac{n\alpha_n[ABD - A^2E - B^2C + A(D^2 - 4EC)n + C(D^2 - 4EC)n^2]}{\alpha_{n-1}[A + (2n+1)C][A + 2nC]}\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (n > 0).
\end{aligned}
$$

By means of these formulas, (5) becomes

$$
\begin{aligned}
&[A + 2nC][A + (2n+1)C][A + (2n+2)C]\, xq_n(x)\\
&\qquad\qquad - [A + (n+1)C][A + 2nC]q_{n+1}(x)\\
(7)\qquad &+ [A + (2n+1)C][AB + AD + (2AD + 2CD)n + 2CDn^2]\, q_n(x)\\
&- n[A + (2n+2)C][ABD - A^2E - B^2C + A(D^2 - 4EC)n\\
&\qquad\qquad + C(D^2 - 4EC)n^2]q_{n-1}(x) = 0.
\end{aligned}
$$

It is easily shown from (2) and (3) that $q_n(x)$ is a polynomial in A, B, C, D, and E as well as x. Therefore (7) is an identity valid for all values of A, B, C, D, E and may be regarded as a recursion formula for *all* sets of polynomials defined by (3).

In order to get the Christoffel-Darboux identity

$$\sum_{k=0}^{n} p_k(t)p_k(x) = \frac{a_n}{a_{n+1}} \frac{p_{n+1}(t)p_n(x) - p_n(t)p_{n+1}(x)}{t - x}$$

for the non-orthogonal sets defined by (3) it is necessary to resort to a sort of analytic extension for the definition of the normalizing factor $1/I_n$. Suppose the constants A, B, C, D, and E are chosen so that $A + nC \neq 0$ for $n > 1$, and so that $ABD - A^2E - B^2C + A(D^2 - 4EC)n + C(D^2 - 4EC)n^2 \neq 0$ for $n > 0$. Then $\alpha_n \neq 0$, and, if I_0 is arbitrarily assigned a value different from zero, the ratio I_n^2/I_{n-1}^2 given by formula (6) serves to define $I_1, I_2, \cdots$ when the integral definition of the I's fails.[3] Writing $p_n(x) = q_n(x)/I_n$, one can return (5) to the form of (4) from which the Christoffel-Darboux identity follows.

3. **Convergence of expansions in the polynomials of the Jacobi type.** When either $\alpha \leqq -1$ or $\beta \leqq -1$, the polynomials

$$\frac{1}{\rho(x)} \frac{d^n}{dx^n} [M^n(x)\rho(x)] = \frac{1}{(x-a)^\alpha (b-x)^\beta} \frac{d^n}{dx^n} [(x-a)^{\alpha+n}(b-x)^{\beta+n}]$$

are not orthogonal for the reason that $\rho(x)$ is not integrable over (a, b). Nevertheless, suitable functions may be expanded in terms of such polynomials.

[3] This definition may require I_n to be imaginary.

To prove this an identity relating the functions $\sum_{k=0}^{n} p_k(t)p_k(x)$ for different weight functions in the case of ordinary orthogonal polynomials is extended to the non-orthogonal polynomials.

On the interval (a, b), let the set of polynomials $[q_n(x)]$ be orthonormal with respect to the weight function $\rho(x)$ and let the polynomials $[p_n(x)]$ be orthonormal with respect to the weight function $\pi_m(x)\rho(x)$, where $\pi_m(x)$ is a polynomial of degree m, non-negative on (a, b). The two systems of polynomials are connected by the identities

$$\pi_m(x)p_n(x) = A_{n0}q_n(x) + A_{n1}q_{n+1}(x) \quad + \cdots + A_{nm}q_{n+m}(x),\text{[4]} \tag{8}$$

$$q_n(x) = A_{n0}p_n(x) + A_{n-1,1}p_{n-1}(x) + \cdots + A_{n-m,m}p_{n-m}(x), \tag{9}$$

where $A_{nk} = \int_a^b \pi_m(x)\rho(x)p_n(x)q_{n+k}(x)\,dx$. By means of these identities one can show that the difference

$$\pi_m(t)\sum_{k=0}^{n} p_k(t)p_k(x) - \sum_{k=0}^{n} q_k(t)q_k(x) \tag{10}$$

is a combination of the p's, whose nature is too involved to warrant precise statement, but is sufficiently well described for present purposes by remarking that it consists of $\frac{1}{2}m(m+1)^2$ terms, each of the form $A_{ij}A_{kl}p_i(x)p_k(t)$, and that $p_{n-m+1}(t)$ is the polynomial of lowest order to appear.

Consider the two systems of polynomials

$$Q_0(x) = 1, \quad Q_n(x) = \frac{1}{(x-a)^\alpha(b-x)^\beta}\frac{d^n}{dx^n}[(x-a)^{\alpha+n}(b-x)^{\beta+n}] \qquad (n = 1, 2, \cdots), \tag{11}$$

$$P_0(x) = 1, \quad P_n(x) = \frac{1}{(x-a)^{\alpha+h}(b-x)^{\beta+k}}\frac{d^n}{dx^n}[(x-a)^{\alpha+h+n}(b-x)^{\beta+k+n}] \qquad (n = 1, 2, \cdots), \tag{12}$$

where α and β are not both > -1, and where h and k are integers such that $\alpha + h > -1, \beta + k > -1$. Let $p_n(x)$ be the normalized system of polynomials corresponding to (12). System (11) cannot be normalized in the ordinary sense, but, with the method of the preceding section and with the exceptions noted, its members can be given quasi-normalized forms. The exceptional cases to be avoided here in defining the I's are most conveniently indicated by rewriting ratio (6)

$$\frac{I_n^2}{I_{n-1}^2} = \frac{-n\alpha_n(b-a)^2(\alpha+n)(\beta+n)}{\alpha_{n-1}(\alpha+\beta+2n)(\alpha+\beta+2n+1)}. \tag{13}$$

[4] Cf. J. Shohat, *On the development of continuous functions in series of Tchebycheff polynomials*, Trans. Amer. Math. Soc., vol. 27(1925), pp. 537–550; p. 542.

Were α and $\beta > -1$, the ratio of the normalizing factors for the polynomials of zero degree of systems (11) and (12) would be

$$\left\{\frac{\int_a^b (x-a)^{\alpha+h}(b-x)^{\beta+k}\,dx}{\int_a^b (x-a)^{\alpha}(b-x)^{\beta}\,dx}\right\}^{\frac{1}{2}} \tag{14}$$
$$= \left\{\frac{(b-a)^{h+k}(\alpha+h)(\alpha+h-1)\cdots(\alpha+1)(\beta+k)(\beta+k-1)\cdots(\beta+1)}{(\alpha+\beta+h+k+1)(\alpha+\beta+h+k)\cdots(\alpha+\beta+2)}\right\}^{\frac{1}{2}}.$$

Now if I_0, to which an arbitrary value different from zero may be assigned in the definition of the I's by ratio (13), is given one of the two values required by relation (14) when the undefined integral in the denominator is replaced by I_0^2, and if $[q_n(x)]$ represents the quasi-normalized system of polynomials so formed from system (11), then the polynomials of the two systems $[p_n(x)]$ and $[q_n(x)]$ are connected by identities (8) and (9) with $\pi_m(t) = (t-a)^h(b-t)^k$. These identities then enable one to express the difference (10) in the form described.

Suppose $f(x)$ is such a function that $(x-a)^{\alpha}(b-x)^{\beta}f(x)$ and $(x-a)^{\alpha-h}\cdot(b-x)^{\beta-k}[f(x)]^2$ are integrable over the interval (a, b). Let

$$S_n(x) = \sum_{i=0}^{n} p_i(x) \int_a^b (t-a)^{\alpha+h}(b-t)^{\beta+k}f(t)p_i(t)\,dt,$$

$$s_n(x) = \sum_{i=0}^{n} q_i(x) \int_a^b (t-a)^{\alpha}(b-t)^{\beta}f(t)q_i(t)\,dt.$$

Then, on multiplying the difference (10) by $(t-a)^{\alpha}(b-t)^{\beta}f(t)$ and integrating over the interval (a, b), one has

$$S_n(x) - s_n(x) = \sum A_{ij}A_{rs}p_i(x)\int_a^b (t-a)^{\alpha+h}(b-t)^{\beta+k}\frac{f(t)}{(t-a)^h(b-t)^k}\,p_r(t)\,dt. \tag{15}$$

It is not difficult to show that the A's are bounded, and it is well known that the p's are bounded at any point interior to (a, b). Since $(x-a)^{\alpha+h}(b-x)^{\beta+k}\cdot[f(x)]^2/[(x-a)^{2h}(b-x)^{2k}]$ is assumed to be integrable, the $\frac{1}{2}m(m+1)^2$ integrals in the second member of (15) tend to zero as n increases indefinitely.[5] This proves

THEOREM I. *If $[p_n(x)]$ is the system of normalized Jacobi polynomials for the weight function $(x-a)^{\alpha+h}(b-x)^{\beta+k}$, where $\alpha+h>-1$, $\beta+k>-1$, $\alpha, \beta \neq -1, -2, \cdots$, and $\alpha+\beta \neq -2, -3, \cdots$, if $[q_n(x)]$ is the system of quasi-normalized polynomials of the Jacobi type for the weight function $(x-a)^{\alpha}(b-x)^{\beta}$, and if $(x-a)^{\alpha}(b-x)^{\beta}f(x)$ and $(x-a)^{\alpha-h}(b-x)^{\beta-k}[f(x)]^2$ are integrable over the interval (a, b), then, when x is an interior point of (a, b), $\lim_{n\to\infty}[S_n(x) - s_n(x)] = 0$.*

[5] For a somewhat similar attack upon the problem of the equivalence of expansions of a function in terms of ordinary orthonormal polynomials see G. H. Peebles, *An equivalence theorem for series of orthogonal polynomials*, Proceedings of the National Academy of Sciences, vol. 25(1939), pp. 97–104.

4. **Polynomials orthogonal with respect to a weight function which changes sign.** The identities leading to Theorem I suggest, in a way which is perhaps not entirely apparent without more detailed explanation than is warranted here, a method of constructing polynomials orthonormal with respect to a weight function which changes sign in the interval of orthogonality. Since $\int_a^b \rho(x)[q_n(x)]^2\,dx$ is not necessarily of the same sign for all values of n, when $\rho(x)$ changes sign in the interval (a, b), the definition of normalization is broadened so that $q_n(x)$ is said to be normalized if $\int_a^b \rho(x)[q_n(x)]^2\,dx = \pm 1$. This definition is used from here on.

It is important to note that each of the polynomials orthogonal and normalized with respect to the weight function $\rho(x)$ has degree corresponding to position in the sequence when $\rho(x)$ changes sign as well as when $\rho(x)$ is non-negative on (a, b).[6] If this were not true, the polynomials could not be normalized. Certainly $q_0(x) \neq 0$. Suppose $q_n(x)$ is the first polynomial to be of lower degree than n. Then $q_n(x) = C_{n-1}q_{n-1}(x) + \cdots + C_0q_0(x)$, if the C's are properly chosen, and as a consequence,

$$\int_a^b \rho(x)[q_n(x)]^2\,dx = \int_a^b \rho(x)q_n(x)[C_{n-1}q_{n-1}(x) + \cdots + C_0q_0(x)]\,dx = 0.$$

Let $[q_n(x) = c_nx^n + \cdots]$ be a set of polynomials orthonormal with respect to the weight function $\rho(x)$ over the interval (a, b). Let

$$Q_{nm}(y_1^{(n_1)}, y_2^{(n_2)}, \cdots, y_k^{(n_k)}) \qquad (n_1 + 1 + n_2 + 1 + \cdots + n_k + 1 = n - m + 1)$$

represent the determinant

$$\begin{vmatrix} q_n(y_1) & q_{n-1}(y_1) & \cdots & q_m(y_1) \\ q_n'(y_1) & q_{n-1}'(y_1) & \cdots & q_m'(y_1) \\ q_n''(y_1) & q_{n-1}''(y_1) & \cdots & q_m''(y_1) \\ \cdots & \cdots & \cdots & \cdots \\ q_n^{(n_1)}(y_1) & q_{n-1}^{(n_1)}(y_1) & \cdots & q_m^{(n_1)}(y_1) \\ q_n(y_2) & q_{n-1}(y_2) & \cdots & q_m(y_2) \\ q_n'(y_2) & q_{n-1}'(y_2) & \cdots & q_m'(y_2) \\ \cdots & \cdots & \cdots & \cdots \\ q_n^{(n_2)}(y_2) & q_{n-1}^{(n_2)}(y_2) & \cdots & q_m^{(n_2)}(y_2) \\ \cdots & \cdots & \cdots & \cdots \\ q_n(y_k) & q_{n-1}(y_k) & \cdots & q_m(y_k) \\ \cdots & \cdots & \cdots & \cdots \\ q_n^{(n_k)}(y_k) & q_{n-1}^{(n_k)}(y_k) & \cdots & q_m^{(n_k)}(y_k) \end{vmatrix},$$

where the y's are distinct.[7]

[6] In consequence, identities (8), (9), and (10) hold when $\rho(x)$ and $\pi_m(x)$ change sign in the interval of orthogonality, provided the polynomials $[q_n(x)]$ and $[p_n(x)]$ exist for all n.

[7] If one of the indices n_i is zero, it is to be dropped from the notation altogether. Thus,

$$Q_{n,n-1}(y_1^{(0)}, y_2^{(0)}) \equiv Q_{n,n-1}(y_1, y_2) = \begin{vmatrix} q_n(y_1) & q_{n-1}(y_1) \\ q_n(y_2) & q_{n-1}(y_2) \end{vmatrix}.$$

The numbers $m, n_1, n_2, \cdots, n_k$ being given, the points for which the determinant $Q_{n+m-1,n}(r_1^{(n_1)}, r_2^{(n_2)}, \cdots, r_k^{(n_k)})$ vanishes for some value of n form a set of measure zero in the k-dimensional space of the arguments $r_1, r_2, \cdots, r_k$. Hence, the r's may be selected from any interval of positive measure, in particular, the interval (a, b), in such a way that $Q_{n+m-1,n}(r_1^{(n_1)}, r_2^{(n_2)}, \cdots, r_k^{(n_k)})$ vanishes for no value of n. Suppose the r's have been selected in this way. Then for all values of n, $Q_{n+m,n}(x, r_1^{(n_1)}, r_2^{(n_2)}, \cdots, r_k^{(n_k)})$ has r_i for a root of order $n_i + 1$ $(i = 1, 2, 3, \cdots, k)$. Let a set of polynomials $[p_n(x)]$ be defined by the relation $\pi_m(x)p_n(x) = \lambda_n Q_{n+m,n}(x, r_1^{(n_1)}, r_2^{(n_2)}, \cdots, r_k^{(n_k)})$,[8] where $\pi_m(x) = (x - r_1)^{n_1+1}(x - r_2)^{n_2+1} \cdots (x - r_k)^{n_k+1}$ and λ_n is an arbitrary constant not zero. One has

$$\int_a^b \rho(x)\pi_m(x)p_n(x)p_j(x)\,dx$$

$$= \begin{cases} 0, & \text{when } n \neq j, \\ (-1)^m\lambda_n^2 \dfrac{c_{n+m}}{c_n} Q_{n+m-1,n}(r_1^{(n_1)}, r_2^{(n_2)}, \cdots, r_k^{(n_k)}) Q_{n+m,\,n+1}(r_1^{(n_1)}, r_2^{(n_2)}, \cdots, r_k^{(n_k)}) \neq 0, & \\ & \text{when } n = j. \end{cases}$$

Hence, the polynomials $[p_n(x)]$ are orthogonal with respect to $\pi_m(x)\rho(x)$ and can be normalized by a proper choice of λ_n.

When the polynomials $[q_n(x)]$ exist for all n, not only is the non-vanishing of $Q_{n+m-1,n}(r_1^{(n_1)}, r_2^{(n_2)}, \cdots, r_k^{(n_k)})$ sufficient, but it is also necessary[9] for the existence of the polynomials $[p_n(x)]$. Three lemmas are useful in the proof.[10]

LEMMA I. *Always*

$$(x - r_1)^{n_1+1}(x - r_2)^{n_2+1} \cdots (x - r_k)^{n_k+1} = \lambda Q_{n0}(x, r_1^{(n_1)}, r_2^{(n_2)}, \cdots, r_k^{(n_k)}),$$

where λ is a suitable constant.

The proof proceeds by the method of mathematical induction. The lemma is obviously true for

$$Q_{1,0}(x, r_1) = \begin{vmatrix} q_1(x) & q_0(x) \\ q_1(r_1) & q_0(r_1) \end{vmatrix} = \frac{1}{\lambda}(x - r_1).$$

Assume that $Q_{n0}(x, r_1^{(n_1)}, r_2^{(n_2)}, \cdots, r_k^{(n_k)})$ is a polynomial of the n-th degree possessing a root r_1 of order $n_1 + 1$, a root r_2 of order $n_2 + 1$, etc., $n_1 + 1 + n_2 +$

[8] This special form of identity (8) is given by Szegö, *Ueber die Entwickelung einer analytischen Funktion nach den Polynomen eines Orthogonalsystems*, Mathematische Annalen, vol. 82(1921), pp. 188–212, footnote on pp. 190–191; Bernstein, *Sur les polynomes orthogonaux relatifs à un segment fini*, Journal de Mathématiques pures et appliquées, (9), vol. 9(1930), pp. 127–177; pp. 138–139.

[9] See references of footnote 2.

[10] So far as the proof of these lemmas is concerned, the property of orthogonality of the q's can be dropped and the q's taken to be any system of polynomials for which the coefficient of x^n in $q_n(x)$ differs from zero.

$1 + \cdots + n_k + 1 = n$. Consider the determinants $Q_{n+1,0}(x, r_1^{(n_1+1)}, r_2^{(n_2)}, \cdots, r_k^{(n_k)})$ and $Q_{n+1,0}(x, r_1^{(n_1)}, r_2^{(n_2)}, \cdots, r_k^{(n_k)}, r_{k+1})$, $r_{k+1} \neq r_1, r_2, \cdots, r_k$. Both are polynomials of the $(n + 1)$-th degree having the roots required by the lemma if the minors of $q_{n+1}(x)$ do not vanish. The respective minors are $Q_{n0}(r_1^{(n_1+1)}, r_2^{(n_2)}, \cdots, r_k^{(n_k)})$ and $Q_{n0}(r_1^{(n_1)}, r_2^{(n_2)}, \cdots, r_k^{(n_k)}, r_{k+1})$, and neither can vanish since $Q_{n0}(x, r_1^{(n_1)}, r_2^{(n_2)}, \cdots, r_k^{(n_k)})$ does not have r_{k+1} for a root nor r_1 for a root of order $n_1 + 2$.

LEMMA II. *If the minors of $q_{m-1}(x), q_{m-2}(x), \cdots, q_0(x)$ in $Q_{n0}(x, r_1^{(n_1)}, r_2^{(n_2)}, \cdots, r_k^{(n_k)})$ all vanish, that is to say, if Q_{n0} is a linear combination of the polynomials $q_n(x), q_{n-1}(x), \cdots, q_m(x)$, then Q_{n0} is a constant multiple of a determinant obtained from Q_{n0} by striking out the last m columns and some, though not necessarily an arbitrary, set of m rows. Any m rows, not including the first, may be struck out, provided the minor of $q_n(x)$ in the final determinant is not zero.*

The matrix obtained from Q_{n0} by striking out the first row, the first column, and the last m columns must be of rank $n - m$. Otherwise Q_{n0} is not of the n-th degree, as can be seen by using the determinants from this matrix to expand Q_{n0} by Laplace's development. Therefore, it is possible to obtain, in the manner stated by the lemma, a determinant $D(x)$ in which the minor of $q_n(x)$ is not zero. By hypothesis,

$$Q_{n0}(x, r_1^{(n_1)}, r_2^{(n_2)}, \cdots, r_k^{(n_k)}) = C_n q_n(x) + C_{n-1} q_{n-1}(x) + \cdots + C_m q_m(x),$$

where C_n is the cofactor of $q_n(x)$, etc. Keeping in mind this form of Q_{n0} and the points where Q_{n0} and its derivatives vanish, one sees that the $n - m + 1$ sets of n constants

$$q_j(r_1), q_j'(r_1), \cdots, q_j^{(n_1)}(r_1), q_j(r_2), \cdots, q_j^{(n_2)}(r_2), \cdots, q_j^{(n_k)}(r_k) \qquad (j = m, m + 1, \cdots, n)$$

taken from the columns of Q_{n0} are linearly dependent with multipliers C_n, $C_{n-1}, \cdots, C_m$. Hence, every determinant of order $n - m + 1$ formed from their matrix is zero. But $D^{(\nu)}(r_i)$, where i has any of the values $1, 2, \cdots, k$ and $0 \leqq \nu \leqq n_i$, is one of these determinants or else has two rows identical. Therefore $D(x)$ has all the roots of Q_{n0} to the same orders.

LEMMA III. *If $Q_{n0}(x, r_1^{(n_1)}, \cdots, r_k^{(n_k)}) = C_n q_n(x) + C_{n-1} q_{n-1}(x) + \cdots + C_m q_m(x)$, but there is a set of indices $n_1', n_2', \cdots, n_i'$ such that $Q_{n0}(x, r_1^{(n_1)}, r_2^{(n_2)}, \cdots, r_k^{(n_k)}) \neq \lambda' Q_{nm}(x, r_1^{(n_1')}, r_2^{(n_2')}, \cdots, r_i^{(n_i')})$, where $n_1' \leqq n_1$, $n_2' \leqq n_2$, etc., $i \leqq k$, then there exist two or more polynomials of the n-th degree which are linearly independent, which have the root r_1 to the order $n_1' + 1$, the root r_2 to the order $n_2' + 1, \cdots$, the root r_i to the order $n_i' + 1$, and which are expressible linearly in terms of the polynomials $q_n(x), q_{n-1}(x), \cdots, q_m(x)$.*

If $Q_{n0} \neq \lambda' Q_{nm}$, then, by Lemma II, $Q_{n-1,m}(r_1^{(n_1')}, r_2^{(n_2')}, \cdots, r_i^{(n_i')}) = 0$ and $Q_{nm}(x, r_1^{(n_1')}, \cdots, r_i^{(n_i')})$ is not a polynomial of the n-th degree. Nevertheless,

the reasoning of Lemma II can be used to show that Q_{nm} has all the roots of Q_{n0} to the same orders. Hence, $Q_{nm} \equiv 0$. But

$$Q_{nm}(x, r_1^{(n_1')}, r_2^{(n_2')}, \cdots, r_i^{(n_i')}) = B_n q_n(x) + B_{n-1}q_{n-1}(x) + \cdots + B_m q_m(x),$$

where B_n is the cofactor of $q_n(x)$ in Q_{nm}, etc. By hypothesis $B_n = Q_{n-1,m}(r_1^{(n_1')}, r_2^{(n_2')}, \cdots, r_i^{(n_i')}) = 0$. Hence, the coefficient of x^{n-1} in the second member is $B_{n-1}C_{n-1}$. Since the coefficient of x^{n-1} must be zero if $Q_{nm} \equiv 0$, and since $C_{n-1} \neq 0$, $B_{n-1} = 0$. Continuation of this procedure shows that $B_n = B_{n-1} = \cdots = B_m = 0$. Let the smallest possible number of rows, say j rows, be replaced in Q_{nm} to get one of the determinants $D(x)$ which, by Lemma II, always exists. If only $j - 1$ rows of the last $n - m$ rows of Q_{nm} are replaced, the resultant determinant, like Q_{nm}, is such that all the cofactors of $q_n(x)$, $q_{n-1}(x)$, $\cdots$, $q_m(x)$ are zero. Hence, any determinant which contains $n - m - j + 1$ rows from the last $n - m$ rows of Q_{nm} is zero. Let any row of $D(x)$ which is not contained in Q_{nm} be replaced by a set of $n - m + 1$ independent parameters and let this determinant be $E(x)$. If it can be shown that $E(x)$ has the root r_1 to the order $n_1' + 1$, etc., then the lemma is proved. Let r_α be one of the roots in question. When $\nu \leqq n_\alpha'$, $E^{(\nu)}(r_\alpha) = 0$. For, if the row $q_n^{(\nu)}(r_\alpha)q_{n-1}^{(\nu)}(r_\alpha) \cdots q_m^{(\nu)}(r_\alpha)$ has not been replaced, and so is one of the rows of $E(x)$, then two rows of $E^{(\nu)}(r_\alpha)$ are identical and, if this row has been replaced, then $E^{(\nu)}(r_\alpha)$ has $n - m - j + 1$ rows from the last $n - m$ rows of Q_{nm}.

The proof that the non-vanishing of $Q_{n+m-1,n}(r_1^{(n_1)}, r_2^{(n_2)}, \cdots, r_k^{(n_k)})$ is necessary for the existence of the p's follows. Assume that the polynomials $[p_n(x)]$ orthonormal with respect to $\pi_m(x)\rho(x)$ exist. Then the p's and q's are connected by identity (9). Suppose $\pi_m(x)p_n(x) \not\equiv \lambda_n Q_{n+m,n}(x, r_1^{(n_1)}, r_2^{(n_2)}, \cdots, r_k^{(n_k)})$. According to Lemma I this cannot occur for $n = 0$. If it occurs for some positive value of n, then by Lemma III there is at least one other polynomial, say $\pi_m(x)\bar{p}_n(x)$, not proportional to $\pi_m(x)p_n(x)$, which is expressible in terms of $q_{n+m}(x), \cdots, q_n(x)$. But the existence of a relation like identity (9) for $\bar{p}_n(x)$ is sufficient to prove that $\bar{p}_n(x)$ is orthogonal to every polynomial of lower degree with respect to $\pi_m(x)\rho(x)$. Therefore, if a_n and $\bar{a}_n$ are respectively the leading coefficients of $p_n(x)$ and $\bar{p}_n(x)$, the G's in the relation

$$\frac{\bar{a}_n}{a_n} p_n(x) - \bar{p}_n(x) = G_{n-1}p_{n-1}(x) + \cdots + G_0 p_0(x)$$

are all zero; for

$$G_j = \int_a^b \pi_m(x)\rho(x)\left[\frac{\bar{a}_n}{a_n} p_n(x) - \bar{p}_n(x)\right]p_j(x)\,dx = 0 \qquad (j = 0, 1, \cdots, n-1).$$

Hence $\bar{a}_n p_n(x) - a_n\bar{p}_n(x) \equiv 0$, and one is led to the contradiction that $p_n(x)$ and $\bar{p}_n(x)$ are linearly dependent. It follows that, if the p's exist, $Q_{n+m-1,n}(r_1^{(n_1)}, r_2^{(n_2)}, \cdots, r_k^{(n_k)}) \neq 0$ for $n = 0, 1, 2, \cdots$.

The results so far obtained are summarized in

THEOREM II. *If the polynomials $[q_n(x)]$ orthonormal with respect to $\rho(x)$ exist for all n, then a necessary and sufficient condition that the polynomials orthonormal with respect to $\pi_m(x)\rho(x)$ exist for all n is that $Q_{n+m-1,n}(r_1^{(n_1)}, r_2^{(n_2)}, \cdots, r_k^{(n_k)}) \neq 0$ for $n = 0, 1, 2, \cdots$.*

Since the p's surely exist if $\rho(x)$ is non-negative and if $r_1, r_2, \cdots, r_k$ lie outside or at the ends of the interval of orthogonality, Theorem II has a

COROLLARY. *If the polynomials $[q_n(x)]$ are orthonormal with respect to a weight function which is non-negative and positive on a set of positive measure in the interval of orthogonality, then $Q_{n+m-1,n}(r_1^{(n_1)}, r_2^{(n_2)}, \cdots, r_k^{(n_k)})$ may vanish only when one or more of the r's is an interior point of the interval of orthogonality or when two or more of the r's have the same value.*

A theorem of orthogonal polynomials states that if $q_n(x)$ is a polynomial from a set of polynomials orthogonal with respect to a non-negative weight function, then $q_n(x)$ has n simple real roots which are interior points of the interval of orthogonality. An extension of this theorem is

THEOREM III. *If the polynomials $[q_n(x)]$ are orthogonal with respect to a non-negative weight function $\rho(x)$, then the Wronskian*

$$\begin{vmatrix} q_{n+m}(x) & q_{n+m-1}(x) & \cdots & q_n(x) \\ q'_{n+m}(x) & q'_{n+m-1}(x) & \cdots & q'_n(x) \\ \cdots & \cdots & \cdots & \cdots \\ q^{(m)}_{n+m}(x) & q^{(m)}_{n+m-1}(x) & \cdots & q^{(m)}_n(x) \end{vmatrix}$$

has no real roots, if m is odd, and has n and only n real roots all of which are simple and are interior points of the interval of orthogonality if m is even.

Since $(x - r)^{m+1}\rho(x)$ is non-negative, if m is odd, the polynomials orthogonal and normalized with respect to $(x - r)^{m+1}\rho(x)$ surely exist. It follows from Theorem II that the Wronskian $Q_{n+m,n}(r^{(m)})$ vanishes for no value of n. Thus, part of the theorem is easily proved.

Suppose m is even and that $Q_{n+m,n}(r^{(m)}) = 0$. By the corollary to Theorem II, r must be an interior point of the interval of orthogonality. Since

$$\frac{d^m}{dx^m} Q_{n+m,n}(x, r^{(m-1)}) \bigg|_{x=r} = Q_{n+m,n}(r^{(m)}) = 0,$$

$Q_{n+m,n}(x, r^{(m-1)})$ has a root of order $m + 1$ at $x = r$. The polynomial $p_n(x)$ corresponding to the weight function $(x - r)^m\rho(x)$ and defined by the relation $(x - r)^m p_n(x) = \lambda_n Q_{n+m,n}(x, r^{(m-1)})$ has a root at r. Conversely, whenever $p_n(x)$ has a root at r, $Q_{n+m,n}(r^{(m)}) = 0$.

The roots of $p_n(x)$, say $y_1, y_2, \cdots, y_n$, are real and distinct and contained in (a, b), since the weight function $(x - r)^m\rho(x)$ is non-negative, and are continuous functions of r; so the function $r - y_i(r)$, which is negative when $r = a$ and positive when $r = b$, vanishes at least once in the interval (a, b). There-

fore $Q_{n+m,n}(r^{(m)})$ has at least n roots in (a, b). To show that $r - y_i(r)$ does not vanish more than once, let x be defined as a multiple-valued function of r by the equation $p_n(x) = 0$. On each branch of the function x represents one of the roots $y_1, y_2, \cdots, y_n$. Since

$$p_n(x) = \frac{\lambda_n Q_{n+m,n}(x, r^{(m-1)})}{(x-r)^m} = \lambda_n \begin{vmatrix} \phi_n(x) & \phi_{n-1}(x) & \cdots & \phi_{n-m}(x) \\ q_{n+m}(r) & q_{n+m-1}(r) & \cdots & q_n(r) \\ \cdots & \cdots & \cdots & \cdots \\ q_{n+m}^{(m-1)}(r) & q_{n+m-1}^{(m-1)}(r) & \cdots & q_n^{(m-1)}(r) \end{vmatrix},$$

where $\phi_i(x)$ is the integral part of the quotient $q_{i+m}(x)/(x - r)^m$, one has

$$\frac{dx}{dr} = -\frac{\begin{vmatrix} \frac{\partial\phi_n}{\partial r} & \frac{\partial\phi_{n-1}}{\partial r} & \cdots & \frac{\partial\phi_{n-m}}{\partial r} \\ q_{n+m}(r) & q_{n+m-1}(r) & \cdots & q_n(r) \\ \cdots & \cdots & \cdots & \cdots \\ q_{n+m}^{(m-1)}(r) & q_{n+m-1}^{(m-1)}(r) & \cdots & q_n^{(m-1)}(r) \end{vmatrix} + \begin{vmatrix} \phi_n & \phi_{n-1} & \cdots & \phi_{n-m} \\ q_{n+m}(r) & q_{n+m-1}(r) & \cdots & q_n(r) \\ \cdots & \cdots & \cdots & \cdots \\ q_{n+m}^{(m-2)}(r) & q_{n+m-1}^{(m-2)}(r) & \cdots & q_n^{(m-2)}(r) \\ q_{n+m}^{(m)}(r) & q_{n+m-1}^{(m)}(r) & \cdots & q_n^{(m)}(r) \end{vmatrix}}{\begin{vmatrix} \frac{\partial\phi_n}{\partial x} & \frac{\partial\phi_{n-1}}{\partial x} & \cdots & \frac{\partial\phi_{n-m}}{\partial x} \\ q_{n+m}(r) & q_{n+m-1}(r) & \cdots & q_n(r) \\ \cdots & \cdots & \cdots & \cdots \\ q_{n+m}^{(m-1)}(r) & q_{n+m-1}^{(m-1)}(r) & \cdots & q_n^{(m-1)}(r) \end{vmatrix}}.$$

The relations

$$\phi_k(x) = \frac{q_{k+m}^{(m)}(r)}{m!} + \frac{q_{k+m}^{(m+1)}(r)}{(m+1)!}(x-r) + \cdots + \frac{q_{k+m}^{(k+m)}(r)}{(k+m)!}(x-r)^k,$$

$$\phi_k(r) = \frac{q_{k+m}^{(m)}(r)}{m!},$$

$$\left.\frac{\partial\phi_k}{\partial r}\right|_{x=r} = \frac{q_{k+m}^{(m+1)}(r)}{m!} - \frac{q_{k+m}^{(m+1)}(r)}{(m+1)!} = m\,\frac{q_{k+m}^{(m+1)}(r)}{(m+1)!},$$

$$\left.\frac{\partial\phi_k}{\partial x}\right|_{x=r} = \frac{q_{k+m}^{(m+1)}(r)}{(m+1)!}$$

show that always

$$\left.\frac{dx}{dr}\right|_{r=x} = -m.$$

Hence

$$\frac{d}{dr}[r - y_i(r)]_{r=y_i} = 1 + m.$$

The function $r - y_i(r)$, being continuous, vanishes but once and so $Q_{n+m,n}(r^{(m)})$ has n and only n real roots. The roots are simple; for, if $Q_{n+m,n}(r^{(m)}) = 0$ and therefore $p_n(r) = 0$,

$$\frac{d}{dr} Q_{n+m,n}(r^{(m)}) = \frac{d^{m+1}}{dx^{m+1}} Q_{n+m,n}(x, r^{(m-1)}) \bigg|_{x=r} = \frac{1}{\lambda_n} \frac{d^{m+1}}{dx^{m+1}} (x - r)^m p_n(x) \bigg|_{x=r} \neq 0.$$

UNIVERSITY OF MINNESOTA.

THEORY OF COGROUPS

By J. E. Eaton

1. Introduction. Grouplike systems with non-unique multiplication have been the subject of several recent papers. In 1938 Dresher and Ore[1] undertook an axiomatic investigation of such systems, which they called multigroups. Some of their most interesting results were concerned with the relation of submultigroups of a multigroup to the multigroup itself. However, for these theorems they found it necessary to restrict their consideration to submultigroups which satisfied a "reversibility" condition.

In this paper we shall examine some of the properties of a special type of multigroup in which every submultigroup is reversible. We have called this particular kind of multigroup a *cogroup* (or, more properly, *left cogroup*). Since the multiplicative system of the left coset decomposition[2] of any group with respect to a subgroup is a cogroup, a few of the results contained in this paper may be of some interest from a group theoretical viewpoint. However, if it can be shown that any cogroup may be generated by the left coset decomposition of a group, many of our theorems would reduce to trivialities. Such a proof, nevertheless, would be of considerable importance. It would permit a formulation of the problem of the extension of groups by non-normal subgroups analogous to the so-called solution of Schreier's for the normal case.

2. Axioms. A *cogroup* (or, more properly, *left cogroup*) is an algebraic system in which there is defined a single binary operation, multiplication, subject to six axioms.

Axiom 1. The Product. *If c_i and c_j are any two elements of a cogroup $\mathfrak{C}$, then the product c_ic_j is a non-void subset of $\mathfrak{C}$.*

$$c_ic_j = \{c_k'\}.$$

The existence of the product of any two elements of $\mathfrak{C}$ permits us to give meaning to the notion of the product of any two subsets of $\mathfrak{C}$. If A and B are two non-void subsets of $\mathfrak{C}$ with elements $\{a_i\}$ and $\{b_i\}$ respectively, then an element c of $\mathfrak{C}$ is in AB if and only if c is contained in some product a_jb_k.

Axiom 2. The Associative Law. *If c_i, c_j, c_k are any three elements of $\mathfrak{C}$, then*

$$(c_ic_j)c_k = c_i(c_jc_k).$$

Received July 18, 1939.

[1] Dresher and Ore, *Theory of multigroups*, American Journal of Mathematics, vol. 60(1938), pp. 705-733.

[2] We shall call a *left coset* of a subgroup $\mathfrak{H}$ a complex of the form $\mathfrak{H}g$.

These products have meaning under our definition of products of subsets.

Axiom 3. The Unit. *There exists an element e in $\mathfrak{C}$ such that*

$$ec_i = c_i$$

for all c_i in $\mathfrak{C}$.

Axiom 4. The Right Inverse. *To each c_i in $\mathfrak{C}$ corresponds a subset C_i of $\mathfrak{C}$ such that*

(i) *if $x \subset C_i$, then $c_i x \supset e$;*

(ii) *if $c_j c_i \supset c_k$, then $c_k C_i \supset c_j$.*

Axiom 5. The Left Inverse. *To each c_i in $\mathfrak{C}$ corresponds at least one element c_i^{-1} such that*

$$c_i^{-1} c_i \supset e.$$

Axiom 6. *If $c_i c_j$ contains k distinct elements, then $c_i c$ contains k distinct elements for any c in $\mathfrak{C}$.*

Although it is not essential in the proof of all of our theorems, we shall throughout this paper assume that $\mathfrak{C}$ contains but a finite number of elements. If, in the above axioms, we interchange the order of multiplication, the postulates would then define what we might call a *right cogroup* and the theorems in this paper would be equally applicable to the amended system.

As an example of a cogroup consider the system whose multiplication scheme is given by the table below.

	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b,c	b,c	e,a	e,a
c	b,c	b,c	e,a	e,a

3. **General properties.** It is quite easy to show that a cogroup is a multigroup. The known properties of multigroups may then be used to simplify the proofs of our theorems.

Theorem 1. *The element e is a right unit:*

$$ce \supset c$$

for any c in $\mathfrak{C}$.

Proof. Let $c_i e \supset c_j$. Then $c_i e \supset c_j e$. Furthermore, $c_j e \supset c_i$ by Axiom 4 (since e is its own only right inverse) and $c_j e \supset c_i e$. Hence $c_i e = c_j e$ and $c_i e \supset c_i$.

As an immediate consequence of this theorem we have that $\mathfrak{C}$ is a multigroup. That is, for any ordered pair of elements c_i, c_j there is at least one x and one y in $\mathfrak{C}$ such that

$$c_i x \supset c_j \quad \text{and} \quad y c_i \supset c_j.$$

THEOREM 2. *If c_1c contains k elements, there exist $c_2, c_3, \cdots, c_k$ (all distinct and different from c_1) such that $c_1c = c_2c = \cdots = c_kc$ for all c in $\mathfrak{C}$.*

Proof. By Axiom 6, c_1e contains k distinct elements $c_1, c_2, \cdots, c_k$. But in the proof of Theorem 1 we showed that $c_ie = c_je$ $(i, j = 1, 2, \cdots, k)$. Hence $c_iec = c_jec$ and $c_ic = c_jc$.

We shall refer to the elements $c_1, c_2, \cdots, c_k$ as the *e-conjugates* of c_1. It is obvious that e-conjugate is a symmetric, reflexive, transitive relation.

THEOREM 3. *If $c_ic \supset c'$ and $c_jc \supset c'$, then c_i and c_j are e-conjugate.*

Proof. Divide $\mathfrak{C}$ into disjoint subsets $C_1, C_2, \cdots$ of e-conjugate elements. Let C_i contain k_i elements and $\mathfrak{C}$ k elements. Then $\sum k_i = k$. But if the theorem were false we would have that $\mathfrak{C}c$ contains less than k elements, and this is impossible since $\mathfrak{C}$ is a multigroup.

4. **Inverses.** The conditions in Axiom 4, enunciated for right inverses, are also satisfied by left inverses.

THEOREM 4. *If $c_ic_j \supset c_k$, then for some left inverse c_i^{-1}, we have $c_i^{-1}c_k \supset c_j$.*

Proof. We may find x such that $xc_k \supset c_j$. Therefore $xc_ic_j \supset c_j$. Then by Theorem 3, $xc_i \supset e$.

The inverses of a cogroup have many properties in common with the inverses of a group. However, the notion of e-conjugate is important in their consideration.

THEOREM 5. *A left inverse is also a right inverse and conversely. The inverses of a given element are e-conjugate and all e-conjugates of an inverse are inverses. The number of e-conjugates of an inverse of a given element is the same as the number of e-conjugates of the given element.*

Proof. Let $c_ic_j \supset e$. Then $c_jc_ic_j \supset c_j$ and by Theorem 3, $c_jc_i \supset e$. If $c_ic_j \supset e$ and $c_kc_j \supset e$, then by Theorem 3, c_i and c_k are e-conjugate. If $c_ic_j \supset e$ then $c_kc_j \supset e$ for any c_k e-conjugate to c_i. Let $c_ic_i^{-1} \supset e$ and let the product contain k elements. Let c_i^{-1} have k' e-conjugates and suppose $k' < k$. Then for any c the number of solutions of $c_ix \supset c$ is less than or equal to k' since we must have $c_i^{-1}c \supset x$. Hence the number of elements in $c_i\mathfrak{C}$ (if multiplicity is regarded) is less than or equal to nk', where n is the number of elements in $\mathfrak{C}$. But the number is nk. Thus k' is not less than k. Similarly k is not less than k' and hence $k = k'$.

We have in the proof of this theorem established the following result.

THEOREM 6. *If c_i has k e-conjugates, then $c_ix \supset c$ has exactly k solutions for any c.*

5. **Subcogroups.** An interesting property of finite groups is that closure under multiplication is sufficient to insure that a subset of a group is itself a group. That characteristic is equally valid in cogroups.

THEOREM 7. *If $\mathfrak{H}$ is a subset of a cogroup $\mathfrak{C}$ for which $\mathfrak{H} \supset \mathfrak{H}\mathfrak{H}$, then $\mathfrak{H}$ is itself a cogroup which is both left and right closed in $\mathfrak{C}$.*[3]

Proof. Let $\mathfrak{H}$ contain k elements. Then for any h in $\mathfrak{H}$ we have $\mathfrak{H} \supset \mathfrak{H}h$. By Theorem 3, $\mathfrak{H}h$ contains at least k elements. Hence $\mathfrak{H} = \mathfrak{H}h$ for all h. Then for any h we may find an h' such that $h'h \supset h$. By Theorem 3, $h' = e$ and e is in $\mathfrak{H}$. Suppose $xh_i \supset h_j$. But we may find an h' such that $h'h_i \supset h_j$. Then by Theorem 3, $h'e \supset x$ and x is in $\mathfrak{H}$. Suppose $h_ix \supset h_j$. Then $\mathfrak{H}x \supset \mathfrak{H}$. Since $\mathfrak{H}x$ contains but k distinct elements, we have $\mathfrak{H}x = \mathfrak{H}$. Then $\mathfrak{H} \supset ex = x$. Thus $\mathfrak{H}$ is both left and right closed and the postulates of a cogroup are readily seen to be satisfied.

The concept of *reversibility* was first introduced by Dresher and Ore.[4] For a submultigroup to generate a coset decomposition of the multigroup, it is sufficient that the submultigroup be reversible. It is thus a highly desirable property for a submultigroup to possess. But the results of Dresher and Ore[5] yield immediately:

THEOREM 8. *Any subcogroup is both left and right reversible.*

We thus have that if $\mathfrak{C}$ is a cogroup and $\mathfrak{H}$ a subcogroup, then a coset decomposition of $\mathfrak{C}$ is possible,[6]

$$\mathfrak{C} = \mathfrak{H} + \mathfrak{H}c_2 + \cdots + \mathfrak{H}c_r,$$

possessing the usual properties of such decompositions, namely: every element of $\mathfrak{C}$ lies in some coset, the cosets are disjoint, each coset contains its generating element, and any element in a coset generates that coset. Since we have already seen that every coset contains the same number of elements, we may state the theorem of Lagrange for cogroups.

THEOREM 9. *The order of a subcogroup (i.e., the number of elements in it) divides the order of the cogroup.*

The problem of determining the conditions under which two elements lie in the same coset is not easy. We may, however, readily establish a result in that direction.

THEOREM 10. *If $\mathfrak{H}$ is a subcogroup, then $\mathfrak{H}c_ie \supset \mathfrak{H}c_j$ if and only if $c_ic_j^{-1} \supset h$, where h is some element of $\mathfrak{H}$.*

6. **Homomorphisms.** It is clear that any group is also a cogroup. It then follows from the theorem below that the multiplicative system of the left coset decomposition of a group with respect to any subgroup is itself a cogroup. Hence all the theorems deduced in this paper are directly applicable to such multiplicative systems.

[3] A submultigroup $\mathfrak{H}$ is *right closed* if every solution x of $hx \supset h'$ lies in $\mathfrak{H}$ when h and h' lie in $\mathfrak{H}$.

[4] Op. cit., p. 715. A submultigroup $\mathfrak{H}$ is *left reversible* if whenever $\mathfrak{H}x \supset y$ then $\mathfrak{H}y \supset x$.

[5] Op. cit., Theorem 6, p. 717.

[6] Dresher and Ore, op. cit., pp. 717–718.

THEOREM 11. *The multiplicative system, $\mathfrak{C}/\mathfrak{H}$, of the left coset decomposition of any cogroup $\mathfrak{C}$ with respect to any subcogroup $\mathfrak{H}$ is a cogroup.*

Proof. Let $\mathfrak{C} = \mathfrak{H} + \mathfrak{H}c_2 + \cdots + \mathfrak{H}c_r$. The various $\mathfrak{H}c_i$ form a multigroup with a left scalar unit $\mathfrak{H}$. Hence Axioms 1, 2, 3, 5 are satisfied. Suppose $\mathfrak{H}c_i\mathfrak{H}c_j \supset \mathfrak{H}c_k$. We may choose c_i and c_j so that $c_ic_j \supset c_k$. Then there is a c_j^{-1} such that $c_kc_j^{-1} \supset c_i$. Then $\mathfrak{H}c_k\mathfrak{H}c_j^{-1} \supset \mathfrak{H}c_i$ where $\mathfrak{H}c_j\mathfrak{H}c_j^{-1} \supset \mathfrak{H}$. Thus Axiom 4 is satisfied. Let $\mathfrak{H}c\mathfrak{H} = \mathfrak{H}c_1e + \mathfrak{H}c_2e + \cdots + \mathfrak{H}c_se$. Then $\mathfrak{H}c\mathfrak{H}x = \mathfrak{H}c_1x + \mathfrak{H}c_2x + \cdots + \mathfrak{H}c_sx$. There are the same number of elements in $\mathfrak{H}c_ie$ and $\mathfrak{H}c_ix$, for $\mathfrak{H}c_ie$ contains all the e-conjugates of any element in it. Then there are the same number of distinct cosets in $\mathfrak{H}c_ie$ and $\mathfrak{H}c_ix$. Hence the number of distinct cosets in $\mathfrak{H}c\mathfrak{H}x$ is less than or equal to the number in $\mathfrak{H}c\mathfrak{H}$. But if $\mathfrak{H}c\mathfrak{H} = \mathfrak{H}c_1 + \mathfrak{H}c_2 + \cdots + \mathfrak{H}c_t$, then $\mathfrak{H}c_1\mathfrak{H} = \mathfrak{H}c_2\mathfrak{H} = \cdots = \mathfrak{H}c_t\mathfrak{H}$ since they each equal $\mathfrak{H}c\mathfrak{H}$ (two double cosets which are not disjoint are equal[7]). Divide $\mathfrak{C}/\mathfrak{H}$ into disjoint subsets $C_1, C_2, \cdots$ of e-conjugate elements and select from each C_i a single representative $\mathfrak{H}c_i$. Let C_i contain k_i members. Then $\sum k_i = r$, where r is the order of $\mathfrak{C}/\mathfrak{H}$. Suppose $\mathfrak{H}c_i\mathfrak{H}x$ contained less than k_i elements. Then $[\sum \mathfrak{H}c_i]\mathfrak{H}x$ contains less than r elements. But $[\sum \mathfrak{H}c_i]\mathfrak{H}x = [\mathfrak{C}/\mathfrak{H}]\mathfrak{H}x$ and, since $\mathfrak{C}/\mathfrak{H}$ is a multigroup, $[\mathfrak{C}/\mathfrak{H}]\mathfrak{H}x$ contains r elements. Hence we have a contradiction and Axiom 6 is thus satisfied.

THEOREM 12. *If a cogroup $\mathfrak{C}$ is homomorphic to a cogroup $\mathfrak{C}^*$, then $\mathfrak{C}^*$ is isomorphic to the coset decomposition of $\mathfrak{C}$ with respect to a subcogroup.*

Proof. Suppose c and c' have the same image c^*. We may find an x in $\mathfrak{C}$ such that $xc \supset c'$. Then $x^*c^* \supset c^*$. By Theorem 3, $x^* = e^*$. The theorem then follows immediately from the theory of multigroups.[8]

7. Normal subcogroups.

In the usual theory of multigroups it is necessary to distinguish between several types of normal submultigroups. The weakest form of normality that has been introduced is left normality. A submultigroup $\mathfrak{H}$ is *left normal*[9] if $\mathfrak{H}m \supset m\mathfrak{H}$ for every m in the multigroup. Even for this extremely weak normality the Jordan-Hölder theorem for composition series is valid. Dresher and Ore define a *strongly normal* submultigroup $\mathfrak{H}$ to be such that $m\mathfrak{H}m^{-1} = \mathfrak{H}$ for every m and every inverse m^{-1}.[10] They show that the necessary and sufficient condition that the quotient multigroup $\mathfrak{M}/\mathfrak{H}$ be a group is that $\mathfrak{H}$ be strongly normal in $\mathfrak{M}$.[11] The strongly normal submultigroups form a Dedekind structure. Thus, there exists a unique minimal strongly normal submultigroup. The situation for cogroups is much simpler.

THEOREM 13. *A left normal subcogroup $\mathfrak{H}$ is strongly normal.*

[7] Dresher and Ore, op. cit., p. 718.

[8] Eaton and Ore, *Remarks on multigroups*, American Journal of Mathematics, vol. 62(1940), Theorem 3, p. 68.

[9] The terminology is mine. Krasner, I believe, calls such a submultigroup *right seminormal*.

[10] Op. cit., Theorem 10, p. 730.

[11] Op. cit., Theorem 12, p. 730.

Proof. $\mathfrak{H}c\mathfrak{H} \subset \mathfrak{H}c$ for any c. Since there is but one coset on the right, $\mathfrak{H}c\mathfrak{H} = \mathfrak{H}c$. Thus every element in $\mathfrak{C}/\mathfrak{H}$ is its own only e-conjugate and $\mathfrak{C}/\mathfrak{H}$ is a group. Hence $\mathfrak{H}$ is strongly normal.

In a recent paper[12] we introduced the notions of conjugate submultigroups and the normalizer of a submultigroup, analogous to those constructs in groups. While it is possible to simplify slightly the defining conditions for conjugate subcogroups, it does not appear easy to construct all the conjugates of a given subcogroup. The *normalizer* of a subcogroup $\mathfrak{H}$, however, may be redefined in a much simpler fashion, namely: the totality of elements c in $\mathfrak{C}$ for which $\mathfrak{H}c \supset c\mathfrak{H}$. It is obvious that the normalizer constitutes the left scalar elements[13] of $\mathfrak{C}/\mathfrak{H}$. That the normalizer is itself a subcogroup is an immediate consequence of:

Theorem 14. *The set of all left scalar elements of a cogroup form a subcogroup.*

The theorem follows directly from the fact that the product of two left scalar elements is a left scalar element.

8. Application to groups.

We may rephrase some of our previous results in a form more closely related to group theory.

Theorem 15. *The necessary and sufficient condition that a partition of a group $\mathfrak{G}$ into a finite number of disjoint subsets be the left coset decomposition of $\mathfrak{G}$ with respect to a subgroup $\mathfrak{H}$ is that the multiplicative system $\mathfrak{M}$ of the partition be such that*

(i) *$\mathfrak{M}$ contains a left scalar unit;*

(ii) *if a product AX_1 in $\mathfrak{M}$ contains k distinct elements, then AX contains k distinct elements for all X in $\mathfrak{M}$.*

It may be of interest to examine a proof of the necessity by purely group theoretical methods.

Let $\mathfrak{G} = \mathfrak{H} + \mathfrak{H}g_2 + \cdots + \mathfrak{H}g_r$ be a left coset decomposition of $\mathfrak{G}$ with respect to $\mathfrak{H}$ and let the cosets $\mathfrak{H}g_i$ be elements of a multiplicative system $\mathfrak{M}$. Then $\mathfrak{H}$ is obviously a left scalar unit of $\mathfrak{M}$. Let $\mathfrak{H}g_i\mathfrak{H}g_j = \mathfrak{H}g_1' + \mathfrak{H}g_2' + \cdots + \mathfrak{H}g_k'$, where the $\mathfrak{H}g_i'$ are distinct. Then $g_i'(g_j')^{-1} \subset \mathfrak{G} - \mathfrak{H}$ for $i \neq j$. Consider any coset $\mathfrak{H}g_s$ and let $g_s = g_jg''$. We then have $\mathfrak{H}g_i\mathfrak{H}g_s = \mathfrak{H}g_i\mathfrak{H}g_jg'' = \mathfrak{H}g_1'g'' + \mathfrak{H}g_2'g'' + \cdots + \mathfrak{H}g_k'g''$. But $g_i'g''(g_j'g'')^{-1} = g_i'(g_j')^{-1} \subset \mathfrak{G} - \mathfrak{H}$ so the k cosets are distinct.

As we have seen, the coset decomposition of a group also possesses the following property.

(iii) *If a product X_1A in $\mathfrak{M}$ contains k distinct elements, then there exist $X_2, X_3, \cdots, X_k$ (all distinct and different from X_1) such that $X_iY = X_jY$ for all Y in $\mathfrak{M}$ and $i, j = 1, 2, \cdots, k$.*

[12] J. E. Eaton, *Associative multiplicative systems*, American Journal of Mathematics, vol. 62(1940), pp. 222-232.

[13] An element c is a *left scalar* if cx is a single element for all x.

A group theoretical proof of this may also be of interest.

Consider for any coset $\mathfrak{H}g_i$ the group $(g_i^{-1}\mathfrak{H}g_i, \mathfrak{H}) = \mathfrak{K}$. Expand $\mathfrak{H}$ in cosets with respect to $\mathfrak{K}$: $\mathfrak{H} = \mathfrak{K} + \mathfrak{K}h_2 + \cdots + \mathfrak{K}h_s$. Then $\mathfrak{H}g_i\mathfrak{H} = \mathfrak{H}g_i + \mathfrak{H}g_ih_2 + \cdots + \mathfrak{H}g_ih_s$, for note first that the cosets are distinct. If $j \neq t$, then $g_ih_th_j^{-1}g_i^{-1} \subset \mathfrak{G} - \mathfrak{H}$ since $h_th_j^{-1} \subset \mathfrak{H} - \mathfrak{K}$. Furthermore, any $\mathfrak{H}g_ih = \mathfrak{H}g_ih_j$ for some j since there exists some h_j such that $h_jh^{-1} \subset \mathfrak{K}$. Hence $\mathfrak{H}g_i\mathfrak{H}g_j = \mathfrak{H}g_ih_2\mathfrak{H}g_j = \cdots = \mathfrak{H}g_ih_s\mathfrak{H}g_j$ for any coset $\mathfrak{H}g_j$, where s is the number of cosets in the product $\mathfrak{H}g_i\mathfrak{H}g_j$.

HOFSTRA COLLEGE.

SOME PROPERTIES OF ${}_3F_2(-n, n+1, \zeta; 1, p; v)$

By S. O. Rice

The fact that the Legendre function $P_n(x)$ may be expressed as the hypergeometric series

$$P_n(x) = F\left(-n, n+1; 1; \frac{1-x}{2}\right) = \sum_{r=0}^{\infty} \frac{(-n)_r(n+1)_r}{r!r!}\left(\frac{1-x}{2}\right)^r,$$

where $(\alpha)_r = \alpha(\alpha+1)\cdots(\alpha+r-1)$, together with the fact that the generalized hypergeometric functions, studied by Bateman and Pasternack,

$$F_n(z) = {}_3F_2\left(-n, n+1, \frac{1+z}{2}; 1, 1; 1\right) = \sum_{r=0}^{\infty} \frac{(-n)_r(n+1)_r\left(\frac{1+z}{2}\right)_r}{r!r!r!},$$

$$F_n^m(z) = {}_3F_2\left(-n, n+1; \frac{1+z+m}{2}; 1, m+1; 1\right)$$

$$= \sum_{r=0}^{\infty} \frac{(-n)_r(n+1)_r\left(\frac{1+z+m}{2}\right)_r}{r!r!(m+1)_r}$$

have many interesting properties, suggests that other hypergeometric series in which $(-n)_r(n+1)_r/(r!r!)$ appears as a factor in the general term may also be of interest.

Here we examine one of these series, namely,

$$H_n(\zeta, p, v) = {}_3F_2(-n, n+1, \zeta; 1, p; v)$$

$$= \sum_{r=0}^{\infty} \frac{(-n)_r(n+1)_r(\zeta)_r}{r!r!(p)_r} v^r.$$

When the generalizations given here are compared with the earlier results, only those of Bateman's involving $F_n(z)$ will be mentioned, although in many cases Pasternack [9][1] has obtained relations which, from the standpoint of generality, lie between those given here and those given by Bateman. This omission is made for the sake of brevity and simplicity. In all of the following work we assume that p is not a negative integer, and that n, unless otherwise stated, is a positive integer so that $H_n(\zeta, p, v)$ is a polynomial of degree n; and we shall omit the subscripts 3 and 2 on the hypergeometric function ${}_3F_2$ when it is convenient to do so.

Received July 18, 1939.

[1] Numbers in brackets refer to the bibliography.

The paper is divided into three sections. In §§1, 2 integrals and series which involve $H_n(\zeta, p, v)$ are obtained. In §3 the behavior of $H_n(\zeta, p, v)$ as n becomes large is studied by applying Darboux's method to a generating function which is given in §2.

1. **Integrals involving** $H_n(\zeta, p, v)$. Integrals for $H_n(\zeta, p, v)$ are obtained from the series when certain factors in the general term are expressed as an integral and the order of integration and summation interchanged. This interchange is always permissible since the series for $H_n(\zeta, p, v)$ never has more than $n + 1$ terms. Perhaps the simplest representation is obtained from

$$\frac{(\zeta)_r}{(p)_r} = \frac{\Gamma(p)}{\Gamma(\zeta)\Gamma(p-\zeta)} \int_0^1 t^{\zeta+r-1}(1-t)^{p-\zeta-1}\,dt, \tag{1.1}$$

a relation which leads to

$$H_n(\zeta, p, v) = \frac{\Gamma(p)}{\Gamma(\zeta)\Gamma(p-\zeta)} \int_0^1 t^{\zeta-1}(1-t)^{p-\zeta-1} P_n(1-2vt)\,dt, \tag{1.2}$$

where $R(p) > R(\zeta) > 0$. Other representations which are not subject to the same restrictions may be obtained by using other integrals for $(\zeta)_r/(p)_r$. In this way we may obtain the two expressions

$$H_n(\zeta, p, v) = \frac{\Gamma(p)\Gamma(1+\zeta-p)}{\Gamma(\zeta)} \frac{1}{2\pi i} \int_L t^{-p}(1-t)^{p-\zeta-1} P_n\left(1 - \frac{2v}{t}\right) dt, \tag{1.3}$$

$$H_n(\zeta, p, v) = \frac{\Gamma(p)\Gamma(1-\zeta)}{\Gamma(p-\zeta)} \frac{1}{2\pi i} \int_L t^{-p}(1-t)^{\zeta-1} P_n\left(1 + \frac{2v(1-t)}{t}\right) dt \tag{1.4}$$

subject to $R(\zeta) > 0$ and $R(p - \zeta) > 0$, respectively. L is any contour starting and terminating at infinity that can be deformed into the straight line joining $\frac{1}{2} - i\infty$ and $\frac{1}{2} + i\infty$ without passing over the points $t = 0$ and $t = 1$. Still other integrals may be obtained by taking L to run along the real axis from $t = +1$ to $t = +\infty$ or from $-\infty$ to 0.

The integral

$$H_n(\zeta, p, v) = \frac{1}{2\pi i} \int^{(0+)} F(\zeta, n+1; p; vt)\left(1 - \frac{1}{t}\right)^n \frac{dt}{t} \tag{1.5}$$

is slightly different from the ones above. It holds for all values of ζ, p and v and may be verified by expanding the binomial factor and integrating termwise.

If ζ and p are interchanged in (1.1), the resulting integral may be used to prove

$$P_n(1-2v) = \frac{\Gamma(\zeta)}{\Gamma(p)\Gamma(\zeta-p)} \int_0^1 t^{p-1}(1-t)^{\zeta-p-1} H_n(\zeta, p, vt)\,dt.$$

This equation holds when $R(\zeta) > R(p) > 0$ and is in a sense reciprocal to equation (1.2). Integrals somewhat similar in appearance to the one just above

may be obtained by applying Mellin's inversion formula to an integral obtained from equation (1.2). Setting $x = t/(1-t)$ in that equation leads to

$$(1.6)\quad H_n(\zeta, p, v) = \frac{\Gamma(p)}{\Gamma(\zeta)\Gamma(p-\zeta)} \int_0^\infty x^{\zeta-1}(1+x)^{-p} P_n\left(1 - \frac{2vx}{1+x}\right) dx,$$

the integral converging absolutely for $R(p) > R(\zeta) > 0$. Therefore, by Mellin's inversion formula,[2]

$$(1.7)\quad \Gamma(p)(1+x)^{-p} P_n\left(1 - \frac{2vx}{1+x}\right) = \frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty} \Gamma(\zeta)\Gamma(p-\zeta)H_n(\zeta, p, v)x^{-\zeta}\, d\zeta,$$

where $R(p) > \sigma > 0$. This equation is a generalization of the equation

$$\operatorname{sech} x\, P_n(\tanh x) = \tfrac{1}{2}\int_{-\infty}^{+\infty} e^{ixz} \operatorname{sech} \tfrac{1}{2}\pi z\, F_n(iz)\, dz$$

given by Bateman in [3]. The analogues of equations (1.2) and (1.6) are also given in [3]. The equation

$$(1.8)\qquad \operatorname{cosech} z\, Q_n(\coth z) = \tfrac{1}{4}\pi \int_{-\infty}^{+\infty} e^{ixz} \operatorname{sech}^2 (\tfrac{1}{2}\pi x)\, F_n(ix)\, dx$$

is given in [2]. A slight generalization of (1.8) may be obtained from (1.7) by setting $\sigma = \frac{1}{2}$ and $v = 1$, taking p to be an integer such that $1 \leqq p \leqq n$, multiplying both sides by the reciprocal of

$$(1+x)\left(u - \frac{1-x}{1+x}\right) = (u-1)\left(1 + x\frac{u+1}{u-1}\right),$$

u being real and greater than unity, and integrating x between 0 and ∞. On the left we have the integral

$$\frac{1}{2}\int_{-1}^{+1} \frac{[1 + u - (u-y)]^{p-1} P_n(y)}{u-y}\, dy = (1+u)^{p-1} Q_n(u),$$

where we have set $y = (1-x)/(1+x)$. On the right side the order of integration may be inverted by de la Vallée Poussin's theorem[3] (the integral in (1.7) converges absolutely since $|\Gamma(\zeta)\, \Gamma(p-\zeta)|$ is $O(|\zeta|^{p-1}e^{-\pi|\zeta|})$ and $H_n(\zeta, p, v)$ is $O(\zeta^n)$ as $\zeta \to \sigma \pm i\infty$). This work leads to the result

$$(1.9)\quad \begin{aligned} &(p-1)!\left(\frac{u+1}{2}\right)^p Q_n(u) \\ &\quad = \frac{1}{4\pi i}\int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} \left(\frac{u+1}{u-1}\right)^\zeta \left(\frac{\pi}{\sin \pi\zeta}\right)^2 (1-\zeta)_{p-1} H_n(\zeta, p, 1)\, d\zeta, \end{aligned}$$

where $(1-\zeta)_{p-1} = (1-\zeta)(2-\zeta)\cdots(p-1-\zeta)$.

[2] An account of Mellin's inversion formula is given in the book *Methoden der Math. Phys.*, by Courant and Hilbert, p. 87.

[3] See T. Bromwich, *Theory of Infinite Series*, p. 504.

Another equation of the nature of an integral equation may be obtained from (1.7) when both sides are multiplied by $x^{s-1}(1+x)^{p-q}$, where $0 < R(s) < R(q)$, and integrated with respect to x between 0 and ∞. If s, p, q are so related that a number σ can be chosen which will satisfy both of the inequalities $0 < \sigma < R(p)$, $0 < R(s - \sigma) < R(q - p)$, the order of integration on the right may be inverted to obtain

$$\Gamma(p) \int_0^\infty x^{s-1}(1+x)^{-q} P_n\left(1 - \frac{2vx}{1+x}\right) dx$$
$$= \frac{1}{2\pi i} \int_{\sigma - i\infty}^{\sigma + i\infty} \Gamma(\zeta)\Gamma(p - \zeta) H_n(\zeta, p, v)\, d\zeta \int_0^\infty \frac{x^{s-\zeta-1}}{(1+x)^{q-p}}\, dx$$

or, from (1.6),

$$\frac{\Gamma(q-p)\Gamma(p)\Gamma(s)\Gamma(q-s)}{\Gamma(q)} H_n(s, q, v)$$

(1.10)

$$= \frac{1}{2\pi i} \int_{\sigma - i\infty}^{\sigma + i\infty} \Gamma(\zeta)\Gamma(p - \zeta)\Gamma(s - \zeta)\Gamma(q - p - s + \zeta) H_n(\zeta, p, v)\, d\zeta.$$

Some of the restrictions on s, p, q made in the proof may be relaxed by analytic continuation after we specify that the path of integration is to be curved, if necessary, so that the set of poles of $\Gamma(\zeta)\ \Gamma(q - p - s + \zeta)$ lie to the left and those of $\Gamma(p - \zeta)\ \Gamma(s - \zeta)$ to the right. This implies that s, p, q are such that the required path of integration may be found; i.e., no pole of one set can coincide with a pole of the other set.

2. **Series involving $H_n(\zeta, p, v)$.** All of the series given in this section except (2.10) were suggested by the results given by Bateman and Pasternack. In some cases it is possible to extend the generalization beyond $H_n(\zeta, p, v)$ to the hypergeometric function ${}_3F_2$, and when it is convenient the more general form will be given.

The generalizations of

$$\sum_{n=0}^{\infty} t^n F_n(z) = \frac{1}{1-t} F\left[\frac{z+1}{2}, \frac{1}{2}; 1; -4t(1-t)^{-2}\right],$$

$$\sum_{m=0}^{\infty} \frac{(-t)^m F_n(-2m-1)}{m!} = e^{-t}\,{}_2F_2[-n, n+1; 1, 1; t]$$
$$= e^{-t} Z_n(t)$$

which are given in [3] and [5], respectively, are

(2.1) $$\sum_{n=0}^{\infty} t^n H_n(\zeta, p, v) = \frac{1}{1-t} F[\zeta, \tfrac{1}{2}; p; -4vt(1-t)^{-2}],$$

(2.2) $$\sum_{m=0}^{\infty} \frac{(-t)^m}{m!} F\left[\begin{matrix} a, b, -m; v \\ c, p, \end{matrix}\right] = e^{-t}\,{}_2F_2[a, b; c, p; vt].$$

In (2.1) $|t|$ must be less than the smaller of the roots of $-4vt = (1-t)^2$ and hence can never exceed unity. In the second relation t and v may have any finite values. These results are obtained by assuming t to be so small that the right sides may be expanded in absolutely convergent double series and then rearranging the terms so as to get power series in t. Since the circle of convergence of a power series extends out to the nearest singularity and the ordinary hypergeometric function has a singularity at the point where its argument equals unity, we see that the series on the left in (2.1) converges under the conditions stated. The convergence of the series on the left in (2.2) follows similarly from the fact that e^{-t} and ${}_2F_2$ are integral functions.

Incidentally, the equation (2.2) may be regarded as a special case of the equation

$$\sum_{n=0}^{\infty} \frac{h^n}{n!} A_n(x) = e^{hx} a(h)$$

which defines the Appell polynomials [1] $A_n(x)$ corresponding to the function $a(h)$.

A generalization of the series [3]

$$\sum_{m=0}^{\infty} t^m F_n(-2m-1) = \frac{1}{1-t} P_n\left(\frac{1+t}{1-t}\right)$$

is suggested when the path of integration in equation (1.7) is closed on the left and the residues at the poles of $\Gamma(\zeta)$ evaluated. Other series of this nature are suggested by the analogues of (1.7) obtained from the different representations of $H_n(\zeta, p, v)$ mentioned in connection with equations (1.3) and (1.4). When these series are written down, it is seen possible to generalize them by replacing H_n by the more general ${}_3F_2$. In this way we obtain the following results which are of the same nature as some of those obtained by P. Humbert [7, 8]:

$$\sum_{m=0}^{\infty} \frac{t^m}{m!} (p)_m F\left[\begin{matrix} a, b, -m; v \\ c, p; \end{matrix}\right] = (1-t)^{-p} F\left(a, b; c; -\frac{vt}{1-t}\right), \tag{2.3}$$
$$|t| < \min(1, |1-v|^{-1}),$$

where the inequality is to be read as " $|t|$ is less than the smaller of the numbers 1 and $|1-v|^{-1}$";

$$\sum_{m=0}^{\infty} \frac{t^m}{m!} (p)_m F\left[\begin{matrix} a, b, p+m; v \\ c, p; \end{matrix}\right] = (1-t)^{-p} F\left(a, b; c; \frac{v}{1-t}\right), \tag{2.4}$$
$$|v| < 1, \qquad |t| < \min(1, |1-v|);$$

$$\sum_{m=0}^{\infty} \frac{t^m}{m!} (p)_m F\left[\begin{matrix} a, b, 1-p; v \\ c, 1-p-m; \end{matrix}\right] = (1-t)^{-p} F(a, b; c; v(1-t)), \tag{2.5}$$
$$|v| < 1, \qquad |t| < \min\left(1, \left|1 - \frac{1}{v}\right|\right), \qquad p \neq \text{integer};$$

$$
(2.6) \qquad \begin{gathered} \sum_{m=0}^{\infty} \frac{t^m}{m!} (p)_m F\left[\begin{matrix} a, b, -m; v \\ c, 1-p-m; \end{matrix}\right] = (1-t)^{-p} F(a, b; c; vt), \\ |t| < \min\left(1, \left|\frac{1}{v}\right|\right), \qquad p \neq \text{integer}. \end{gathered}
$$

In these relations $(p)_m = p(p+1) \cdots (p+m-1)$. They may be proved and the regions of convergence established by the procedure used in dealing with equation (2.1). When we set $a = -n$, $b = n+1$, $c = 1$, the ${}_3F_2$ functions may be replaced by H_n and the restriction involving v may be dropped leaving only the conditions $|t| < 1$ and, in the proper places, $p \neq$ integer. For example, (2.3) leads to the equation

$$
(2.7) \qquad \sum_{m=0}^{\infty} \frac{t^m}{m!} (p)_m H_n(-m, p, v) = (1-t)^{-p} P_n\left(1 + \frac{2vt}{1-t}\right), \qquad |t| < 1.
$$

When v is set equal to unity in this equation and in the similar equation obtained for (2.4), we obtain two results given by Bateman [4].

The analogue of Heine's expansion

$$
(2.8) \qquad \frac{1}{s-\mu} = \sum_{n=0}^{\infty} (2n+1) Q_n(s) P_n(\mu),
$$

which is valid when the point μ lies inside the ellipse which passes through the point s and has the point ± 1 for its foci, given in [3] is (the notation has been changed)

$$
\frac{1}{s-1}\left(\frac{s-1}{s+1}\right)^{\frac{1}{2}(z+1)} = \sum_{n=0}^{\infty} (2n+1) Q_n(s) F_n(z).
$$

Probably the simplest method of obtaining the extension of this result is to multiply both sides of the integral (1.2) for $H_n(\zeta, p, v)$ by $(2n+1)\, Q_n(s)$ and sum n from 0 to ∞. If the point $1 - 2v$ lies inside the ellipse passing through s having its foci at ± 1, the series

$$
\sum_{n=0}^{\infty} (2n+1) Q_n(s) P_n(1-2vt)
$$

converges uniformly with respect to t in the interval of integration $0 \leqq t \leqq 1$, and the order of summation and integration may be inverted and the series summed by (2.8). The integral may be expressed as a hypergeometric function and we obtain

$$
(2.9) \qquad \sum_{n=0}^{\infty} (2n+1) Q_n(s) H_n(\zeta, p, v) = \frac{1}{s-1} F\left(\zeta, 1; p; \frac{2v}{1-s}\right).
$$

The method used establishes the result only for $R(p) > R(\zeta) > 0$. When the representations (1.3) and (1.4) are used for $H_n(\zeta, p, v)$, the result is established for $R(\zeta) > 0$ and $R(p - \zeta) > 0$, respectively. This suggests that (2.9)

is true generally, subject only to the conditions that p be not a negative integer and that $1 - 2v$ lie inside the ellipse. This is confirmed by the asymptotic expansions (3.1) and (3.2) for H_n when $0 \leqq v \leqq 1$.

Another series may be obtained by a similar procedure in which

$$e^{ir\mu} = \left(\frac{\pi}{2r}\right)^{\frac{1}{2}} \sum_{n=0}^{\infty} i^n(2n + 1)P_n(\mu)J_{n+\frac{1}{2}}(r)$$

is used in place of Heine's expansion (2.8). It is

$$(2.10) \qquad e^{ir}{}_1F_1(\zeta; p; -2irv) = \left(\frac{\pi}{2r}\right)^{\frac{1}{2}} \sum_{n=0}^{\infty} i^n(2n + 1)H_n(\zeta, p, v)J_{n+\frac{1}{2}}(r)$$

and appears to hold for all values of r, v, ζ and p, although our method of proof requires either $R(\zeta) > 0$ or $R(p - \zeta) > 0$.

Bateman has obtained the expansion[4]

$$(2.11) \qquad \frac{1}{R} P_n\left(\frac{1 + r}{R}\right) P_n\left(\frac{1 - r}{R}\right) = \sum_{m=0}^{\infty} r^m P_m(\cos\theta)\, F_n(-2m - 1)$$

in which n is not necessarily an integer and

$$(2.12) \qquad R^2 = 1 - 2r\cos\theta + r^2.$$

Although up until now we have taken n to be a positive integer, we shall drop this restriction temporarily. Then since (2.7) is derived from (2.3), we see that it holds for general values of n if $|t| < \min(1, |1 - v|^{-1})$, and in it we set $p = 1$ and

$$t = r\frac{\cos\theta - r - i\sin\theta\cos\varphi}{1 - r\cos\theta - ir\sin\theta\cos\varphi}.$$

We shall assume that θ, φ, r, v are real and that r and v lie between 0 and 1. It then follows from the definition of t that $|t| \leqq r$. Thus the series in (2.7) converges uniformly with respect to φ provided both the restrictions $0 < r < 1$ and $|r(1 - v)| < 1$ are satisfied. We may, therefore, integrate the series termwise, after multiplying both sides of the equation by $(1 - t)$, using

$$(2.13) \qquad \begin{aligned} &\frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{(\alpha - \beta\cos\varphi)^{\zeta-1}}{(\gamma - \delta\cos\varphi)^{\zeta}}\, d\varphi \\ &\quad = \frac{1}{(\gamma^2 - \delta^2)^{\frac{1}{2}}}\left(\frac{\alpha^2 - \beta^2}{\gamma^2 - \delta^2}\right)^{\frac{1}{2}(\zeta-1)} P_{\zeta-1}\left(\frac{\alpha\gamma - \beta\delta}{[(\alpha^2 - \beta^2)(\gamma^2 - \delta^2)]^{\frac{1}{2}}}\right), \end{aligned}$$

a relation which holds when the factors of the integrand do not vanish in the interval of integration. The arguments of the various quantities entering the equation may be obtained by analytic continuation from the case in which α, β, γ, δ are real and $0 < \beta < \alpha$, $0 < \delta < \gamma$. In our application of (2.13) $\zeta = m + 1$, $\alpha^2 - \beta^2 = \gamma^2 - \delta^2 = R^2$, and the argument of $P_{\zeta-1}$ is $\cos\theta$.

[4] I am indebted to Professor Bateman for communicating this result to me.

The term remaining on the other side of the equation is the integral of $P_n(1 + 2vt(1-t)^{-1})$. Straightforward substitution shows that

$$1 + \frac{2vt}{1-t} = \frac{1}{R^2}[R'^2 - r^2v^2 - i\,2vr\sin\theta\cos\varphi],$$

where $R'^2 = 1 - 2r(1-v)\cos\theta + r^2(1-v)^2$. Since R, R', rv may be viewed as the three sides of a triangle in which R' always exceeds vr, it follows that

$$a = \frac{R' + vr}{R} \geqq 1, \qquad b = \frac{R' - vr}{R} \leqq 1,$$

where $b = 1$ when $\theta = 0$ and $a = 1$ when $\theta = \pi$. Thus, since

$$R^4(a^2 - 1)(1 - b^2) = 4v^2r^2\sin^2\theta,$$

we may write

$$1 + \frac{2vt}{1-t} = ab - i[(a^2-1)(1-b^2)]^{\frac{1}{2}}\cos\varphi;$$

and from the addition theorem for Legendre functions it follows that the integral of P_n is

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} P_n(ab - i[(a^2-1)(1-b^2)]^{\frac{1}{2}}\cos\varphi)\,d\varphi = P_n(a)P_n(b).$$

Our generalization of (2.11) is therefore

$$\text{(2.14)} \qquad \frac{1}{R}P_n(a)P_n(b) = \sum_{m=0}^{\infty} r^m P_m(\cos\theta)\,H_n(-m, 1, v)$$

in which R is given by equation (2.12) and

$$a = R^{-1}[[1 - 2r(1-v)\cos\theta + r^2(1-v)^2]^{\frac{1}{2}} + vr],$$

$$b = R^{-1}[[1 - 2r(1-v)\cos\theta + r^2(1-v)^2]^{\frac{1}{2}} - vr].$$

The number n is not restricted to integer values, but in the proof θ, r, v have been taken to be real with $0 < r < 1$, $0 < v < 1$. The effect of increasing θ by π in (2.14) is the same as that obtained by changing the sign of r and the restriction $0 < r < 1$ may therefore be replaced by $-1 < r < 1$. Setting $v = 1$ gives the result (2.11).

When we replace $P_m(\cos\theta)$ by Callandreau's integral

$$P_m(\cos\theta) = \frac{1}{m!}\int_0^{\infty} e^{-x\cos\theta}J_0(x\sin\theta)\,x^m\,dx,$$

in (2.14), and use the result, obtained from (2.2),

$$\text{(2.15)} \qquad \sum_{m=0}^{\infty}\frac{(rx)^m}{m!}H_n(-m, 1, v) = e^{rx}{}_2F_2(-n, n+1; 1, 1; -rxv),$$

we obtain a generalization of a result given by Bateman [4]

$$R^{-1}P_n(a)P_n(b) = \int_0^\infty e^{-x(\cos\theta - r)} J_0(x \sin\theta)\, {}_2F_2[-n, n+1; 1, 1; -vrx]\, dx. \tag{2.16}$$

This result may be verified when $|\theta| < \frac{1}{4}\pi$, $0 \leqq r < 2^{\frac{1}{2}} \cos(|\theta| + \frac{1}{4}\pi)$, $0 < v < 1$, and n is a positive integer. Analytic continuation may then be used to show that the equation is also true when θ, r, and v are less restricted; for example, it may be shown to hold when the real part of $\cos\theta - r$ is positive, θ being real and v being finite. One way of verifying equation (2.16) is first to replace ${}_2F_2$ by its series expression, which has $n + 1$ terms, and consider the integral of each term. These integrals may be evaluated by expanding $e^{rx} J_0(x \sin\theta)$ as a power series and integrating termwise. The termwise integration may be shown to be legitimate, when the above conditions are satisfied, by a theorem due to G. H. Hardy.[5] Furthermore, these same conditions guarantee that the resulting multiple series is absolutely convergent and may be rearranged until it takes the form of the series on the right of equation (2.14). Equation (2.16) then follows immediately.

When we set $p = 1$ in (2.9), it becomes

$$\sum_{n=0}^{\infty} (2n+1) Q_n(s) H_n(\zeta, 1, v) = \frac{(s-1)^{\zeta-1}}{(s-1+2v)^{\zeta}}, \tag{2.17}$$

and we know that this certainly holds for all values of ζ since the regions $R(\zeta) > 0$ and $R(1-\zeta) > 0$ overlap. This result may be generalized. We first set

$$s = ab - [(a^2-1)(b^2-1)]^{\frac{1}{2}} \cos\varphi,$$

where $a > b > 1$, so $s > 1$, and assume that $0 < v < 1$ so that the point $1 - 2v$ always lies inside the required ellipse. Since the series converges uniformly with respect to φ, it may be integrated termwise if we use

$$Q_n(a)P_n(b) = \frac{1}{2\pi} \int_{-\pi}^{\pi} Q_n(ab - [(a^2-1)(b^2-1)]^{\frac{1}{2}} \cos\varphi)\, d\varphi.$$

When the right side of (2.17) is integrated with respect to φ between the limits $-\pi$ and π, an integral of the form (2.13) is obtained. Thus we are led to

$$\begin{aligned} &\sum_{n=0}^{\infty} (2n+1) Q_n(a) P_n(b) H_n\left(\zeta, 1, \frac{1-u}{2}\right) \\ &\qquad = S^{-\zeta}(a-b)^{\zeta-1} P_{\zeta-1}\left[\frac{S^2 + (1-u)(u-ab)}{(a-b)S}\right] \end{aligned} \tag{2.18}$$

in which we have replaced v by $\frac{1}{2}(1-u)$ and have placed

$$S^2 = a^2 + b^2 + u^2 - 1 - 2uab.$$

[5] E. C. Titchmarsh, *Theory of Functions*, §1.79.

By means of the asymptotic expansion (3.2) for $H_n(\zeta, 1, v)$ we can show that the restriction $b > 1$ may be removed. Thus equation (2.18) certainly holds for $1 < a$, $-1 < b < a$, $-1 < u < 1$, but these are not necessary conditions as may be seen by setting $u = 1$.

When $u = -1$ this becomes an expansion given in [4], and when $\zeta = 1$ it reduces to a known expansion.

3. **Behavior of** $H_n(\zeta, p, v)$ **for large values of** n. An idea of the behavior of $H_n(\zeta, p, v)$ for large values of u may be obtained by applying Darboux's method [6] of obtaining asymptotic expansions from generating functions. We shall show that

$$H_n(\zeta, p, 1) \sim \frac{\Gamma(p)n^{-2\zeta}}{\Gamma(p-\zeta)\Gamma(1-\zeta)} + (-1)^n \frac{\Gamma(p)n^{2\zeta-2p}}{\Gamma(\zeta-p+1)\Gamma(\zeta)} \tag{3.1}$$

and, if $0 < \varphi < \pi$,

$$\begin{aligned} H_n(\zeta, p, \sin^2 \tfrac{1}{2}\varphi) \sim{}& \frac{\Gamma(p)(n \sin \frac{1}{2}\varphi)^{-2\zeta}}{\Gamma(p-\zeta)\Gamma(1-\zeta)} \\ &+ \frac{\Gamma(p)(n \tan \frac{1}{2}\varphi)^{\zeta-p}}{\Gamma(\zeta)}\left[\frac{2}{n\pi \sin \varphi}\right]^{\frac{1}{2}} \cos\left[(n+\tfrac{1}{2})\varphi + \tfrac{1}{2}\pi(\zeta - p - \tfrac{1}{2})\right]. \end{aligned} \tag{3.2}$$

Each term which appears is to be regarded as the first term of an asymptotic series. In general, one of the terms will be negligible in comparison with the other. In the derivation it is assumed that $\zeta - \frac{1}{2}$ and $p - \zeta - \frac{1}{2}$ are not integers, that $2\zeta - 1$ is not zero or a positive integer. In addition for the $v = 1$ case it is assumed that $2p - 2\zeta - 1$ is not zero or a positive integer. However, the forms of the expression above indicate that these restrictions may not be necessary.

The first term of the asymptotic expansion for $F_n(z)$ may be obtained from (3.1) by setting $p = 1$, $\zeta = \frac{1}{2}(z+1)$. Thus

$$F_n(z) \sim \frac{n^{-z-1}}{[\Gamma(\frac{1}{2}(1-z))]^2} + (-1)^n \frac{n^{z-1}}{[\Gamma(\frac{1}{2}(1+z))]^2}$$

and if $z = iy$ is purely imaginary,

$$F_n(iy) \sim \begin{cases} \dfrac{2}{n\pi} \cosh \tfrac{1}{2}\pi y \cos(2\alpha - y \log n), & n \text{ even}, \\[2ex] \dfrac{2i}{n\pi} \cosh \tfrac{1}{2}\pi y \sin(2\alpha - y \log n), & n \text{ odd}, \end{cases}$$

where $\alpha = \arg \Gamma(\frac{1}{2}(1+iy))$.

The generating function of $H_n(\zeta, p, v)$ is, from equation (2.1),

$$(1-t)^{-1} F[\zeta, \tfrac{1}{2}; p; -4vt(1-t)^{-2}]. \tag{3.3}$$

When considered as a function of t, it will have singularities, in general, at $t = 1$ and at the values of t which make the argument of the hypergeometric function

unity. In order to investigate the behavior of (3.1) near $t = 1$ we may use the relation between hypergeometric functions of argument x and x^{-1}, and the behavior near the other singularities may be obtained by using the relation between hypergeometric functions of argument x and $1 - x$. Thus, retaining only the leading terms, we see that near $t = 1$ the generating function (3.3) assumes the form

$$\frac{\Gamma(p)\Gamma(\frac{1}{2} - \zeta)}{\Gamma(\frac{1}{2})\Gamma(p - \zeta)} \frac{(1 - t)^{2\zeta-1}}{(4v)^{\zeta}} + \frac{\Gamma(p)\Gamma(\zeta - \frac{1}{2})}{\Gamma(\zeta)\Gamma(p - \frac{1}{2})2v^{\frac{1}{2}}}. \tag{3.4}$$

In taking this step we assume that p is not zero or a negative integer and that $\frac{1}{2} - \zeta$ is not an integer. Also since we are concerned only with the singularities of (3.3), we note that we may neglect the second term in (3.4) since it and its associated hypergeometric series have no singularity at $t = 1$. Furthermore, in order that the first term may have a singularity at $t = 1$ we must assume that $2\zeta - 1$ is not zero or a positive integer.

The behavior of (3.3) near the point where $1 = -4vt/(1 - t)^2$ is similarly given by

$$\frac{\Gamma(p)\Gamma(p - \zeta - \frac{1}{2})}{\Gamma(p - \zeta)\Gamma(p - \frac{1}{2})(1 - t)} + \frac{\Gamma(p)\Gamma(\zeta + \frac{1}{2} - p)}{\Gamma(\frac{1}{2})\Gamma(\zeta)} \frac{[(1 - t)^2 + 4vt]^{p-\zeta-\frac{1}{2}}}{(1 - t)^{2p-2\zeta}}, \tag{3.5}$$

where we have assumed that $p - \zeta - \frac{1}{2}$ is not an integer. We may neglect the first term because it has no singularity at the point under consideration ($v \neq 0$). Also, when $v = 1$, $2p - 2\zeta - 1$ must not be zero or a positive integer.

We now consider the case $v = 1$. It is seen that the singularities of the generating function are at $t = 1$ and $t = -1$ both of which lie on the circle of convergence of the series in (2.1). Expressions (3.4) and (3.5) show that the expression

$$\frac{\Gamma(p)\Gamma(\frac{1}{2} - \zeta)}{\Gamma(\frac{1}{2})\Gamma(p - \zeta)} \frac{(1 - t)^{2\zeta-1}}{4^{\zeta}} + \frac{\Gamma(p)\Gamma(\zeta + \frac{1}{2} - p)}{\Gamma(\frac{1}{2})\Gamma(\zeta)} \frac{(1 + t)^{2p-2\zeta-1}}{2^{2p-2\zeta}} \tag{3.6}$$

has the same leading terms as the generating function has at the singularities on the circle of convergence. Hence, by Darboux's method, the first terms in the asymptotic series for $H_n(\zeta, p, 1)$ are given by the coefficient of t^n in the expansion of (3.6) in powers of t. The expression in (3.1) is obtained by considering n to be so large in this coefficient that the relation

$$\frac{\Gamma(\alpha + n)}{\Gamma(n + 1)} \sim n^{\alpha-1}$$

may be used.

The case for $0 < v < 1$ may best be considered by setting $v = \frac{1}{2}(1 - \cos\varphi) = \sin^2 \frac{1}{2}\varphi$, where $0 < \varphi < \pi$, as this enables us to write

$$(1 - t)^2 + 4vt = (e^{i\varphi} - t)(e^{-i\varphi} - t)$$

and the singularities of the generating function lie on the unit circle at $t = 1$, $t = t_1 = e^{i\varphi}$ and $t = t_2 = e^{-i\varphi}$. The expression corresponding to (3.6) is

$$\frac{\Gamma(p)\Gamma(\frac{1}{2} - \zeta)}{\Gamma(\frac{1}{2})\Gamma(p - \zeta)} \frac{(1 - t)^{2\zeta-1}}{(4v)^{\zeta}}$$
$$+ \frac{\Gamma(p)\Gamma(\zeta + \frac{1}{2} - p)}{\Gamma(\frac{1}{2})\Gamma(\zeta)} \left[\frac{(t_2 - t_1)^{p-\zeta-\frac{1}{2}}}{(1 - t_1)^{2p-2\zeta}} (t_1 - t)^{p-\zeta-\frac{1}{2}} + \frac{(t_1 - t_2)^{p-\zeta-\frac{1}{2}}}{(1 - t_2)^{2p-2\zeta}} (t_2 - t)^{p-\zeta-\frac{1}{2}} \right].$$

When this is expanded in a power series in t, the coefficient of t^n gives, after some reduction, expression (3.2).

Bibliography

1. P. Appell, *Sur une classe de polynomes*, Annales Sci. de l'Ecole Normale Supérieure, (2), vol. 9(1880), pp. 119–144.
2. H. Bateman, *The polynomial $F_n(x)$*, Annals of Math., vol. 35(1934), pp. 767–775.
3. H. Bateman, *Some properties of a certain set of polynomials*, Tôhoku Math. Journal, vol. 37(1933), pp. 23–38.
4. H. Bateman, *Spheroidal and bipolar coördinates*, this Journal, vol. 4(1938), pp. 39–50.
5. H. Bateman, *Two systems of polynomials for the solution of Laplace's integral equation*, this Journal, vol. 2(1936), pp. 569–577.
6. G. Darboux, *Mémoire sur l'approximation des fonctions de très-grands nombres*, Journal de Math. pures et appl., (3), vol. 4(1878), pp. 5–56, 377–416.
7. P. Humbert, *Sur des suites de polynomes*, Journal de Math. pures et appl., (9), vol. 4(1925), pp. 355–379.
8. P. Humbert, *Quelques développements hypergéométriques*, Annales Soc. Sci. de Bruxelles, vol. 46(1926), pp. 91–97.
9. S. Pasternack, *A generalization of the polynomial $F_n(x)$*, Philosophical Magazine, (7), vol. 28(1939), pp. 209–226.

Bell Telephone Laboratories.

LA LOI DE JORDAN-HÖLDER DANS LES HYPERGROUPES ET LES SUITES GÉNÉRATRICES DES CORPS DE NOMBRES $\mathfrak{P}$-ADIQUES

Par Marc Krasner

M. Kuntzmann[1] et MM. Ore et Dresher[2] ont montré que, pour les sous-hypergroupes reversibles d'un hypergroupe, a lieu une loi de Jordan-Hölder avec des hypothèses calquées sur celles de la loi de Jordan-Hölder pour les groupes sous sa forme classique. Cette forme de la loi de Jordan-Hölder a l'inconvénient de ne donner dans un cas particulier (celui des hypergroupes de classes), important pour la théorie des corps, que ce que donne déjà la loi de Jordan-Hölder dans les groupes: en effet, si G est un groupe, et si $\bar{G}$ et $g \subset \bar{G}$ sont deux sous-groupes de G, l'hypergroupe quotient droit $\bar{H}$ de $\bar{G}$ par g est invariant dans l'hypergroupe quotient droit H de G par g si, et seulement si, $\bar{G}$ l'est dans G, et l'hypergroupe quotient droit de H par $\bar{H}$ est isomorphe à celui de G par $\bar{G}$.

Je montre dans le travail qui suit que le théorème de Jordan-Hölder a lieu pour les sous-hypergroupes reversibles d'un hypergroupe, sous des hypothèses qui, tout en étant équivalentes aux hypothèses ordinaires dans le cas où l'hypergroupe dont il s'agit est un groupe, sont beaucoup plus faibles que ces dernières dans le cas général.

Il est à remarquer qu'il s'agit dans le présent travail de la forme *stricte* de la loi Jordan-Hölder, et non des analogues, pour les hypergroupes, des formes affaiblies de cette loi,[3] qui ont lieu quand on remplace la condition de l'invariance par celle de la permutabilité ou de la quasi-invariance, etc. La recherche de conditions aussi faibles que possible pour que ces analogues aient lieu semble être un problème intéressant et sans trop grandes difficultés.[4]

Il se trouve que si H et $\bar{H} \subset H$ sont deux sous-hypergroupes de l'hypergroupe de ramification $V_{K/k}$ d'un corps K/k de nombres $\mathfrak{P}$-adiques,[5] et s'il n'existe aucun hypergroupe $\bar{\bar{H}}$ entre $\bar{H}$ et H, $\bar{H}$ possède dans H la propriété qui remplace l'invariance dans la loi de Jordan-Hölder sous ma forme (par contre, $\bar{H}$ n'est invariant dans H que dans des cas exceptionnels). Ce résultat, qui est, en quelque sorte, l'analogue d'un théorème connu de Sylow pour les p-groupes, permet de démontrer un théorème très précis sur la structure des corps $\mathfrak{P}$-

Received July 31, 1939.

[1] *Opérations multiformes. Hypergroupes*, Comptes Rendus, Paris, vol. 204(1937), pp. 1787–1788.

[2] *Theory of multigroups*, American Journal of Mathematics, vol. 60(1938), pp. 705–733.

[3] Voir O. Ore, *On the theorem of Jordan-Hölder*, Trans. Amer. Math. Soc., vol. 41(1937), pp. 266–275, et O. Ore, *Structures and group theory*, I, ce Journal, vol. 3(1937), pp. 149–174.

[4] Certains résultats de M. Marty (*Sur les groupes et hypergroupes attachés à une fraction rationnelle*, Annales de l'Ecole Norm. Sup., vol. 53(1936), pp. 83–123; voir les pp. 96–99) peuvent être regardés comme un essai en ce sens.

[5] Voir la définition de cet hypergroupe dans Krasner, *Sur la primitivité des corps $\mathfrak{P}$-adiques* (Mathematica, vol. 13(1937), pp. 72–191), pp. 83–84.

adiques, que j'avais énoncé dans un de mes travaux précédents.[6] La 2-ème partie du présent travail, consacrée à cette application aux corps $\mathfrak{P}$-adiques, sera publiée plus tard dans ce journal.[7]

Chapitre I

La loi de Jordan-Hölder dans les hypergroupes

1. **Hypergroupes.** Les *hypergroupes* avaient été introduits par M. F. Marty.[8] M. O. Ore[9] les appelle aussi *multigroupes*. Voici leur définition.

Un ensemble H, où est définie une loi de composition de ses éléments, s'appelle hypergroupe par rapport à cette loi si les trois conditions suivantes sont vérifiées:

1. *Le composé ab d'un $b \in H$ par un $a \in H$ est un sous-ensemble jamais vide de H.*

Si A, B sont deux sous-ensembles de H, AB désignera la réunion de tous les ab, $a \in A$, $b \in B$. En particulier si $A = \{a\}$, $B = \{b\}$ $(a, b \in H)$, on écrira aB, Ab, ab au lieu de $\{a\}B$, $A\{b\}$, $\{a\}\{b\}$.

2. $(ab)c = a(bc)$ $(a, b, c \in H)$ (*loi d'associativité*).

3. *Pour tout $a \in H$, $aH = Ha = H$.*

Un sous-ensemble h de H s'appelle *sous-hypergroupe*[10] *de H* s'il est un hypergroupe par rapport à la loi de composition de H. Il est dit *clos*[11] dans H si $(H - h)h = h(H - h) = H - h$. Si h est clos dans H, et que $A \subseteq h$, $B \subseteq H$, on a $A[B \cap (H - h)] \subseteq h(H - h)$, donc $AB \cap h = A(B \cap h)$; de même, dans ce cas, $BA \cap h = (B \cap h) A$.

[6] Voir la p. 125 du travail qu'on vient de citer.

[7] Nous aurons à citer au cours de cette première partie de notre travail les mémoires suivants:

M. Dresher et O. Ore, *Theory of multigroups*, Amer. Jour. of Math., vol. 60(1938), pp. 705–733, qui sera cité comme D.O.

L. Kaloujnine, *Une méthode de construction de sous-groupes infra-invariants*, Paris Comptes Rendus, vol. 208(1939), pp. 1869–1871, qui sera cité comme Ka.

M. Krasner, *Sur la théorie de la ramification des idéaux des corps non-galoisiens de nombres algébriques* (Thèse, Paris), Mémoires (in 4°) de l'Acad. de Belgique (Classe des Sc.), vol. 11(1937), fasc. 4, pp. 1–110; *Sur la primitivité des corps $\mathfrak{P}$-adiques*, Mathematica (Cluj), vol. 13(1937), pp. 72–191; *Une généralisation de la notion de sous-groupe invariant*, Paris Comptes Rendus, vol. 208(1939), pp. 1867–1869, qui seront cités comme K.1, K.2, K.3.

J. Kuntzmann, *Opérations multiformes. Hypergroupes*, Paris Comptes Rendus, vol. 204(1937), pp. 1787–1788, qui sera cité comme Ku.

F. Marty, *Sur une généralisation de la notion de groupe*, Huitième congrès des mathématiciens scandinaves, Stockholm, 1934, pp. 45–49; *Sur les groupes et hypergroupes attachés à une fraction rationnelle*, Annales de l'Ecole Norm. Sup., vol. 53(1936), pp. 83–123, qui seront cités comme M.1 et M.2.

O. Ore, *Structures and group theory*, I, ce Journal, vol. 3(1937), pp. 149–174, qui sera cité comme O.

H. S. Wall, *Hypergroups*, Amer. Jour. of Math., vol. 59(1937), qui sera cité comme W.

[8] M.1; M.2, p. 89.

[9] O., p. 153; D.O., p. 705.

[10] Cette notion est due à M. Marty, voir M.2, p. 93.

[11] Cette notion est due à MM. Dresher et Ore, voir D.O., p. 114. M. Marty emploie le terme "fermé".

Lemme 1. *Si h et $\bar{H} \supseteq h$ sont deux sous-hypergroupes d'un hypergroupe H tels que h soit clos dans $\bar{H}$ et $\bar{H}$ le soit dans H, h est clos dans H.*

Démonstration. On a $(H - h)h = (H - \bar{H})h \cup (\bar{H} - h)h \subseteq (H - \bar{H})\bar{H} \cup (\bar{H} - h)h = (H - \bar{H}) \cup (\bar{H} - h) = H - h$, et, de même $h(H - h) \subseteq H - h$. Mais puisque $Hh = hH = H$ et $hh = h$, on a aussi $(H - h)h \supseteq H - h$ et $h(H - h) \supseteq H - h$. Donc $(H - h)h = h(H - h) = H - h$.

Lemme 2. *Si h et h' sont deux sous-hypergroupes d'un hypergroupe H, dont h est clos dans H, $h \cap h'$ est un sous-hypergroupe clos de h'.*

Démonstration. Soit $\bar{h} = h \cap h'$. On a $(h' - \bar{h})\bar{h} \subseteq (H - h)h \cap h' = (H - h)h' = h' - \bar{h}$ et $\bar{h}(h' - \bar{h}) \subseteq h' - \bar{h}$. Donc, si $c \in \bar{h}$, $c(h' - \bar{h})$ et $(h' - \bar{h})c$ sont disjoints avec $\bar{h}$. Donc puisque $ch' = h'c = h' \supseteq \bar{h}$, on a $c\bar{h} \supseteq \bar{h}$ et $\bar{h}c \supseteq \bar{h}$. Or $\bar{h}\bar{h} \subseteq hh \cap h'h' = h \cap h' = \bar{h}$, d'où $c\bar{h} = \bar{h}c = \bar{h}$ et $(h' - \bar{h})\bar{h} = \bar{h}(h' - \bar{h}) = h' - \bar{h}$, ce qui prouve le lemme.[12]

Un élément e de H tel que, pour tout $c \in H$, on ait $c \in ec$, $c \in ece$, $c \in ce$ s'appelle dans ces cas respectifs *unité gauche,*[13] *centrale, droite*[14] *de H.* Un sous-hypergroupe clos h de H contient toutes les unités gauches et toutes les unités droites de H.[15]

Deux hypergroupes H et H' sont dits *isomorphes*[16] s'il existe une correspondance biunivoque E de H à H' (dite *un isomorphisme de H à H'*) telle que, pour tous $a, b \in H$, $Ea \cdot Eb = E(ab)$ (notation: $H \simeq H'$).

2. Classifications régulières. Hypergroupes quotients. Sous-hypergroupes reversibles. Soit Γ une relation classifiante dans un hypergroupe H. C étant une réunion de classes dans H suivant Γ, C/Γ désignera l'ensemble de toutes les classes suivant Γ contenues dans C.

Γ sera dite une *classification régulière de H* si le composé de deux classes quelconques suivant Γ est une réunion de classes suivant Γ.

Soit Γ une classification régulière et soient C_1, C_2 deux classes suivant Γ. Organisons l'ensemble H/Γ par une loi de composition tel que le composé $(C_1/\Gamma)(C_2/\Gamma)$ d'un élément C_2/Γ de H/Γ par un élément C_1/Γ de H/Γ soit C_1C_2/Γ. Cette composition est associative et $(C_1/\Gamma)(C_2/\Gamma)$ est un sous-ensemble jamais vide de H/Γ. De plus, si $C/\Gamma \in H/\Gamma$ on a $(C/\Gamma)(H/\Gamma) = CH/\Gamma = H/\Gamma$ et $(H/\Gamma)(C/\Gamma) = HC/\Gamma = H/\Gamma$.

Donc, H/Γ ainsi organisé est un hypergroupe qui sera appelé *hypergroupe quotient de H par* Γ.[17]

[12] Un résultat un peu plus restreint que l'ensemble de ces deux lemmes, à savoir que *l'intersection de deux sous-hypergroupes clos d'un hypergroupe en est un sous-hypergroupe clos*, avait été prouvé par MM. Dresher et Ore. Voir D.O., p. 715.

[13] Notion due à MM. Ore et Dresher, D.O., p. 707.

[14] Notion due à MM. Ore et Dresher, D.O., p. 707. Toutefois, la notion des *unités bilatères* est due à M. Marty, M.2, p. 95.

[15] Résultat du à MM. Ore et Dresher, D.O., p. 715.

[16] Notion due à M. Marty, M.2, p. 96.

[17] Comparer aux considérations de K.1, p. 25; la même notion avait été introduite indé-

Soit h un sous-hypergroupe de H, et soit $c \in H$. Les ensembles hc, hch, ch s'appellent respectivement *classe gauche*, *classe centrale* (ou *catégorie*), *classe droite* de c suivant h.

Il se peut que les classes gauches, ou centrales, ou droites suivant h d'éléments de H soient classes au sens de la théorie des ensembles, c'est-à-dire que les classes de deux éléments quelconques c_1, c_2 de H soient ou bien égales ou bien disjointes. D'ailleurs, dans ce cas, tout $c \in H$ est élément de sa classe (respectivement gauche, centrale, droite): en effet, puisque $hH = hHh = Hh = H$, il existe un $c' \in H$ tel que respectivement $c \in hc'$, $c \in hc'h$, $c \in c'h$. Mais alors on a respectivement $hc \subseteq h \cdot hc' = hc'$, $hch \subseteq h \cdot hc'h \cdot h = hc'h$, $ch \subseteq c'h \cdot h = c'h$. Ceci entraîne respectivement $hc = hc'$, $hch = hc'h$, $ch = c'h$; donc respectivement $c \in hc$, $c \in hch$, $c \in ch$.

Si les classes gauches ou centrales ou droites suivant h d'éléments de H sont des classes au sens de la théorie des ensembles, la subdivision de H en ces classes est une relation classifiante dans H qui sera désignée dans ces cas respectifs par $h^{(g)}$, $h^{(c)}$, $h^{(d)}$. *$h^{(\alpha)}$ ($\alpha = g, c, d$) est une classification régulière de H*, parce que si $c, c' \in H$, $hchc'$, $hch \cdot hc'h = hchc'h$, $chc'h$ sont les réunions de toutes les classes respectivement gauches, centrales, droites des éléments de chc'. Les hypergroupes $H/h^{(g)}$, $H/h^{(c)}$, $H/h^{(d)}$ seront appelés (quand ils existent) *hypergroupes quotients gauche, central, droit de H par h.*[18]

Appelons, avec MM. Ore et Dresher,[19] h *reversible* dans H quand et les classes gauches et les classes droites suivant h dans H sont des classes au sens de la théorie des ensembles. Ces auteurs avaient montré[20] que dans ce cas les classes centrales suivant h dans H le sont aussi. Un sous-hypergroupe reversible est toujours clos.

Nous allons énumérer quelques propriétés des sous-hypergroupes reversibles, prouvées par MM. Ore et Dresher, qui nous seront nécessaires dans la suite du travail.

Si h, h' sont deux sous-hypergroupes de H, dont h est reversible et h' est clos dans H, $h \cap h'$ est un sous-hypergroupe reversible de h'.[21]

Si h, h' sont deux sous-hypergroupes reversibles de H, l'intersection de tous les surhypergroupes communs de h et h' est un sous-hypergroupe (h, h') de H, le plus petit surhypergroupe commun de h et de h'. (h, h') est le plus petit surensemble commun A de h et de h' tel que $AA \subseteq A$. (h, h') est reversible dans H.[22]

pendamment par M. Marty dans son cours fait en 1939 à la fondation Claude Peccot (Collège de France).

[18] Ces notions sont dues à MM. Dresher et Ore, D.O., pp. 717–718. Toutefois, dans le cas où H est un hypergroupe *normal* au sens de M. Marty (voir M.2, p. 95) et h est un sous-hypergroupe normal de H, ces notions avaient été introduites par M. Marty, M.2, p. 95 et p. 99.

[19] D.O., p. 715; M. Marty emploie le terme "inversible".

[20] D.O., p. 718.

[21] D.O., p. 716.

[22] D.O., pp. 716–717.

Si h est un sous-hypergroupe reversible de H, et si $\bar{H} \supseteq h$ est une réunion de classes gauches, ou centrales, ou droites suivant h dans H, $\bar{H}/h^{(\alpha)}$ (respectivement $\alpha = g, c, d$) est un sous-hypergroupe de $H/h^{(\alpha)}$ si, et seulement si, $\bar{H}$ est un sous-hypergroupe de H.[23]

Si h et $\bar{H} \supseteq h$ sont deux sous-hypergroupes reversibles de H, $\bar{H}/h^{(\alpha)}$ ($\alpha = g, c, d$) est reversible dans[24] *$H/h^{(\alpha)}$ et $(H/h^{(\alpha)})/(\bar{H}/h^{(\alpha)})^{(\alpha)} \simeq H/\bar{H}^{(\alpha)}$.*[25]

Ce dernier résultat sera completé par une réciproque partielle qui est le

LEMME 3. *Si h est un sous-hypergroupe reversible de H, et si $\bar{H} \supset h$ est un sous-hypergroupe de H tel que $\bar{H}/h^{(g)}$ soit reversible dans $H/h^{(g)}$ et $\bar{H}/h^{(d)}$ le soit dans $H/h^{(d)}$, $\bar{H}$ est reversible dans H.*

Démonstration. Soient c_1, $c_2 \in H$, et posons $c_i^{(g)} = hc_i/h^{(g)}$, $c_i^{(d)} = c_ih/h^{(d)}$ ($i = 1, 2$). On a $(\bar{H}/h^{(g)})c_i^{(g)} = \bar{H}hc_i/h^{(g)} = \bar{H}c_i/h^{(g)}$. Comme $(\bar{H}/h^{(g)})c_1^{(g)}$ et $(\bar{H}/h^{(g)})c_2^{(g)}$ sont disjointes ou coïncident, il en est de même pour $\bar{H}c_1$ et $\bar{H}c_2$. De même, puisque $c_i^{(d)}(\bar{H}/h^{(d)}) = c_ih\bar{H}/h^{(d)} = c_i\bar{H}/h^{(d)}$, $c_1\bar{H}$ et $c_2\bar{H}$ ou bien sont disjointes, et ou bien coïncident.

3. **Hypergroupes$_G$, hypergroupes$_C$, hypergroupes$_D$.** Les groupes sont un cas particulier d'hypergroupes, et les sous-hypergroupes d'un groupe ne sont autre chose que ses sous-groupes. Un groupe G a les propriétés suivantes bien connues:

(a) *Tout sous-groupe g de G y est reversible.*

(b) *Si g et $\bar{G} \supseteq g$ sont deux sous-groupes de G, toutes les classes droites suivant $\bar{G}$ dans G contiennent un même nombre de classes droites distinctes suivant g dans G, et toutes les classes gauches suivant $\bar{G}$ dans G contiennent un même nombre de classes gauches distinctes suivant g dans G. Ces nombres sont égaux tous les deux à l'indice $(\bar{G}:g)$ de g dans $\bar{G}$.*

Quand G est d'ordre fini, on a encore

(c) Un sous-ensemble g de G en est un sous-groupe si, et seulement si, $gg \subseteq g$.

g étant un sous-groupe d'un groupe G, en vertu de (a) tous les quotients $G/g^{(g)}$, $G/g^{(c)}$, $G/g^{(d)}$ sont définis. Un hypergroupe H qui est isomorphe à un hypergroupe de la forme $G/g^{(g)}$, $G/g^{(c)}$, $G/g^{(d)}$ sera appelé dans ces cas respectifs *hypergroupe$_G$*, *hypergroupe$_C$*,[26] *hypergroupe$_D$*.[27]

Si le groupe G dont il est question peut être choisi de manière à être d'ordre fini, H sera dit hypergroupe$_G$, respectivement hypergroupe$_C$, respectivement

[23] D.O., p. 719. Dans des cas particuliers ce résultat avait été prouvé auparavant par MM. Marty (M.2, p. 93, p. 95), Krasner (K.1, p. 26, p. 28), Kuntzmann (Ku., p. 1788), Ore (O., p. 154), Wall.

[24] D.O., p. 719.

[25] Se trouve, d'une manière implicite, dans D.O., p. 719. Démontré auparavant dans des cas particuliers, par MM. Marty (M.2, p. 95), Krasner (K.1, p. 28; K.2, p. 78), Kuntzmann (Ku., p. 1788).

[26] Notion due à M. Marty, M.2, p. 9.

[27] Notion due à l'auteur, K.1, p. 29.

hypergroupe$_D$ *fini.*[28] (a) montre que *tout sous-hypergroupe d'un hypergroupe$_G$, ou d'un hypergroupe$_C$, ou d'un hypergroupe$_D$ est reversible.* (b) montre que *toutes les classes gauches dans un hypergroupe$_G$ H suivant un de ses sous-hypergroupes h ont le même nombre d'éléments égal à celui de h, et il en est de même pour les classes droites dans un hypergroupe$_D$ suivant un de ses sous-hypergroupes. Le nombre d'éléments de $H/h^{(\alpha)}$ ($\alpha = g$ resp. d) est égal au quotient $(H:h)$ du nombre d'éléments de H par celui de h (dit indice de h dans H).*[29]

D'autre part, puisque tous les sous-hypergroupes de G sont les sous-groupes de G, et puisque $g/g^{(\alpha)}$ ($\alpha = g, c, d$) est la seule unité de $G/g^{(\alpha)}$, on voit que *tous les sous-hypergroupes d'un hypergroupe$_\alpha$ ($\alpha = G, C, D$) sont des hypergroupes$_\alpha$; un hypergroupe$_\alpha$ ($\alpha = G, C, D$) H possède une et une seule unité 1_H. Si H est un hypergroupe$_G$, hypergroupe$_C$, hypergroupe$_D$, on a respectivement, pour tout $c \in H$, $1_H c = \{c\}$, $1_H c = cl_H = \{c\}$, $cl_H = \{c\}$. H étant un hypergroupe$_G$, ou un hypergroupe$_C$, ou un hypergroupe$_D$, et h étant un sous-hypergroupe de H, respectivement $H/h^{(g)}$, $H/h^{(c)}$, $H/h^{(d)}$ est encore un hypergroupe de la même nature.*

Si H est un hypergroupe$_\alpha$ fini, (c) *montre qu'un sous-ensemble h de H en est un sous-hypergroupe si, et seulement si, $hh \subseteq h$.*[30]

4. **Chaînes d'hypergroupes. Suites génératrices.** Une suite de sous-hypergroupes d'un hypergroupe H

$$(S) \qquad H = H_0, H_1, H_2, \cdots, H_s = h$$

s'appelle chaîne d'hypergroupes (au sens strict) si, pour tout $i = 1, 2, \cdots, s$, $H_i \subset H_{i-1}$. s est dit *la longueur* de la chaîne (S).

La chaîne (S) est dite *close*, si, pour tout $i = 1, 2, \cdots, s$, H_i est clos dans H_{i-1}, ce qui, en vertu du lemme 1, a lieu si, et seulement si, H_i est clos dans H. La chaîne (S) est dite *reversible* si, pour tout $i = 1, 2, \cdots, s$, H_i est reversible dans H_{i-1}. Une chaîne reversible est aussi close. α étant un des signes g, c, d, et (S) étant une chaîne reversible d'hypergroupes, la suite

$$\frac{H_0}{H_1^{(\alpha)}}, \frac{H_1}{H_2^{(\alpha)}}, \cdots, \frac{H_{s-1}}{H_s^{(\alpha)}}$$

s'appelle *la suite des α-quotients de (S)*. Quand H est un hypergroupe$_G$, ou un hypergroupe$_D$ (auquel cas toute chaîne de la forme (S) est reversible), sa suite de g- respectivement d-quotients sera dite sa *suite de quotients* tout court et il est utile d'introduire encore la *suite d'indices* de (S) qui est la suite $(H_0:H_1)$, $(H_1:H_2)$, $\cdots$, $(H_{s-1}:H_s)$.

La chaîne (S) s'appelle une *suite génératrice de H à partir de h* quand elle est reversible et quand il n'existe aucune chaîne reversible entre H et h autre que

[28] Notion due à l'auteur, K.1, p. 29.
[29] Prouvé par l'auteur dans K.1, p. 27, pour les hypergroupes$_D$.
[30] Dans le cas des hypergroupes$_D$ finis avait été prouvé par l'auteur dans K.1, p. 26.

(S) et dont (S) est une suite partielle. En particulier, si H est un hypergroupe$_G$, ou un hypergroupe$_C$, ou un hypergroupe$_D$, les suites génératrices de H à partir de 1_H seront appelées suites génératrices de H tout court.

5. **Sous-hypergroupes infra-invariants, semi-invariants, invariants. Types. Propriétés des sous-hypergroupes infra-invariants.** Un sous-hypergroupe h de H s'appelle *infra-invariant* dans H si, pour tout $c \in H$, on a $hch \subseteq hc \cup ch$.

Supposons qu'un sous-hypergroupe h de H soit reversible et infra-invariant dans H. Soit $c \in H$. On a $hch = hc \cup ch$. Supposons que $hch \neq ch$. Alors, il existe un c', $c' \in hc$, c' non $\in ch$, tel que $hch \supseteq c'h$. Donc, puisque $c'h$ est disjoint avec ch, on a $hc \supseteq c'h$. Donc $hc = h \cdot hc \supseteq hc'h = hch$. Donc, puisque $hch \supseteq hc$, on a $hch = hc$. Donc *ou bien* $hch = hc$, *ou bien* $hch = ch$.

L'élément c de H sera dit *g-élément*, *c-élément* ou *d-élément* par rapport à h suivant que respectivement $hc = hch \neq ch$, $hc = hch = ch$, $hc \neq hch = ch$.

On a $hc \supset ch$, $hc = ch$, $hc \subset ch$ suivant que c est un g-, c-, ou d-élément par rapport à h.

On désignera par $H_g(h)$, $H_c(h)$, $H_d(h)$ les ensembles des g-, c-, et d-éléments de H par rapport à h. Soit h un sous-hypergroupe reversible infra-invariant de H. La condition nécessaire et suffisante pour que $H/h^{(c)} = H/h^{(g)}$ est que toutes les catégories suivant h soient des classes gauches suivant h, c'est-à-dire que tous les éléments de H en soient des g- ou des c-éléments; c'est-à-dire que, pour tout $c \in H$, on ait $hc \supseteq ch$.

Un tel sous-hypergroupe h de H sera dit *semi-invariant à droite dans H*.[31]

De même, la condition nécessaire et suffisante pour que $H/h^{(c)} = H/h^{(d)}$ est que, pour tout $c \in H$, on ait $hc \subseteq ch$. Les hypergroupes satisfaisant à cette condition seront dits *semi-invariants à gauche dans H*.[31]

h sera dit *semi-invariant dans* H s'il y est ou bien semi-invariant à droite, ou bien semi-invariant à gauche.[32]

La semi-invariance est la condition nécessaire et suffisante pour que $H/h^{(c)}$, $H/h^{(g)}$ et $H/h^{(d)}$ ne soient pas distincts deux à deux.

h sera dit *invariant dans* H s'il y est semi-invariant à gauche et à droite en même temps; c'est-à-dire si, pour tout $c \in H$, on a $hc = ch$. L'invariance de h dans H est la condition nécessaire et suffisante pour que $H/h^{(g)} = H/h^{(d)}$. Et alors $H/h^{(c)} = H/h^{(g)} = H/h^{(d)}$.[33] *Dans un groupe d'ordre fini la notion de l'infra-invariance se confond avec celle de l'invariance*, parce que si g est un sous-groupe d'un groupe fini G, et si $c \in G$, gc et cg ont le même nombre fini d'éléments et aucune des inégalités $gc \supset cg$, $gc \subset cg$ n'est possible. Il n'en est plus de même pour les groupes d'ordre infini: j'ai montré,[34] en effet, qu'il existe de tels

[31] Notion déjà introduite par l'auteur dans K.2, p. 76.

[32] La notion déjà introduite par l'auteur dans le cas des hypergroupes$_D$. Voir K.1, p. 27.

[33] Notion due à Krasner, K.1, p. 24, et, indépendamment à Kuntzmann, dans le cas des sous-hypergroupes normaux d'un hypergroupe normal, Ku., p. 1788.

[34] Voir K.3, p. 1868; je dois aussi signaler la note de M. Leo Kaloujnine (Ka.) dans laquelle il donne une méthode très générale de construction de tels sous-groupes.

groupes G ayant des sous-groupes g infra-invariants, et non invariants. D'ailleurs, un sous-groupe infra-invariant, mais non invariant g de G ne peut pas être semi-invariant, parce que de $gc \supset cg$ résulte $gc^{-1} \subset c^{-1}g$ et inversement, donc aucun de $G_g(g)$, $G_d(g)$ n'est vide.

Dans un hypergroupe$_D$ fini la notion de l'infra-invariance se confond avec celle de la semi-invariance à droite, ce qu'il suffit de prouver pour les quotients $H = G/g^{(d)}$ d'un groupe d'ordre fini G par un de ses sous-groupes g.

En effet supposons qu'il existe un $c \in H$ qui est un d-élément par rapport à un sous-hypergroupe infra-invariant $\bar{H} = \bar{G}/g^{(d)}$ de H.

Soit $c = ag/g^{(d)}$ $(a \in G)$. Alors $\bar{H}c = \bar{G}ag/g^{(d)}$, $c\bar{H} = ag\bar{G}/g^{(d)} = a\bar{G}/g^{(d)}$. En vertu de l'hypothèse faite, on doit avoir $\bar{G}ag \subset a\bar{G}$. Donc $\bar{G}ag$ doit avoir moins d'éléments que $a\bar{G}$, donc que $\bar{G}a \subsetneq \bar{G}ag$, ce qui est absurde.

Correlativement, *dans un hypergroupe$_G$ la notion de l'invariance se confond avec celle de la semi-invariance à gauche.*

Il est aussi à remarquer que si H est un hypergroupe$_\alpha$ $(\alpha = G, C, D)$, $\{1_H\}$ est un sous-hypergroupe infra-invariant de H, parce qu'on a dans ces cas respectifs, pour tout $c \in H$, $\{1_H\}c = \{c\} \subsetneq c\{1_H\}$, $\{1_H\}c = \{c\} = c\{1_H\}$, $\{1_H\}c \supsetneq \{c\} = c\{1_H\}$.

Soient A, B deux sous-ensembles disjoints d'un hypergroupe H. Le couple ordonné (A, B) s'appellera un type dans H. (A, B) et (A', B') étant deux types dans H, (A, B) sera dit *sous-type de* (A', B'), et (A', B') sera dit *sur-type de* (A, B) si $A \subsetneq A'$ et $B \subsetneq B'$. Deux types (A, B), (A', B') dans H seront dit *compatibles* s'il existe un sur-type commun de ces deux types, c'est-à-dire si $A \cap B'$ et $B \cap A'$ sont vides.

$\bar{H}$ étant un sous-hypergroupe de H, $(A \cap \bar{H}, B \cap \bar{H})$ s'appellera *type induit par* (A, B) *dans* $\bar{H}$. Deux types (A, B), (A', B') dans H seront dits *compatibles sur* $\bar{H}$ si les types qu'ils induisent dans $\bar{H}$ sont compatibles.

Un sous-hypergroupe infra-invariant reversible h de H sera dit *du type* (A, B) *sur* H si $H_g(h) \subsetneq A$ et $H_d(h) \subsetneq B$. h sera dit du type (A, B) *au sens strict* sur H, et (A, B) sera dit *type propre de h sur H*, si $H_g(h) \subsetneq A$ et $H_d(h) \subsetneq B$. Deux sous-hypergroupes infra-invariants reversibles h et h' de H seront dits *semblables sur H* si leurs types propres sur H sont compatibles. La relation d'être semblables sur H n'est pas, en général, transitive. On dira d'un sous-hypergroupe $h \subsetneq \bar{H}$ de H infra-invariant et reversible dans $\bar{H}$ qu'il est d'un type (A, B) sur $\bar{H}$, où (A, B) est un type dans H, si h est sur $\bar{H}$ du type induit par (A, B) dans H.

En particulier, 0 désignant l'ensemble vide, on voit que h est semi-invariant à droite, à gauche, invariant dans H si, et seulement si, il est du type respectivement $(H, 0)$, $(0, H)$, $(0, 0)$.

LEMME 4. *Si (A, B) est le type propre sur H d'un sous-hypergroupe h de H, on a $hAh = A$, $hBh = B$.*

Démonstration. Soient c, c' deux éléments de H tels que $hch = hc'h$. Alors $hc' \subsetneq hch$ et $c'h \subsetneq hch$. Puisque h est infra-invariant dans H, une au moins

des égalités $hch = hc$ ou $hch = ch$ doit avoir lieu. Mais dans ces cas respectifs on a $hc' \subseteq hc$ ou $c'h \subseteq ch$, ce qui entraîne respectivement $hc' = hc = hc'h$ ou $c'h = ch = hc'h$. En interchangeant h et h', on voit que l'une des deux égalités $hch = hc$, $hc'h = hc'$ entraîne l'autre et qu'il en est de même pour les deux égalités $hch = ch$, $hc'h = c'h$. Tout est prouvé.

Lemme 5. *h, h' étant deux sous-hypergroupes reversibles et infra-invariants d'un hypergroupe H tel que $H = hh'$, les types propres (A, B), (A', B') de h et de h' sur H sont compatibles si, et seulement si, un des cas suivants a lieu*: (1) *un des h, h' est invariant dans H*; (2) *h et h' sont semi-invariants d'un même côté dans H.*

Démonstration. La compatibilité des (A, B) et (A', B') dans les cas (1), (2) est manifeste. Il s'agit donc de prouver que si (A, B), (A', B') sont compatibles un des cas (1), (2) a lieu. Supposons que $A \cap B' = A' \cap B = 0$. On a, en vertu du lemme 4, $Ah = A$, $Bh = B$, $A'h = A'$, $B'h = B'$. Soit que h n'est pas invariant dans H. Alors ou bien $A \neq 0$, ou bien $B \neq 0$. Si $A \neq 0$, on a $Ah' = Ahh' = AH = H$, donc toute classe suivant $h'^{(d)}$ dans H a des éléments communs avec A. Mais B', quand cet ensemble n'est pas vide, contient au moins une classe suivant $h'^{(d)}$, ce qui entraîne $A \cap B' \neq 0$, contre l'hypothèse faite. Donc $B' = 0$. De même, si $B \neq 0$, on doit avoir $A' = 0$. Donc, à moins que l'on ait $A' = B' = 0$, on doit avoir $A = A' = 0$ ou $B = B' = 0$, ce qui prouve la proposition.

Théorème 1. *Soient h et h' deux sous-hypergroupes reversibles, infra-invariants et semblables de H, et soit (h, h') leur hypergroupe composé, c'est-à-dire le plus petit sous-hypergroupe de H qui les contient. Alors (h, h') est encore infra-invariant dans H et est semblable à h et à h'. En particulier, si h et h' sont semi-invariants à droite, ou à gauche, ou invariants dans H, (h, h') l'est aussi.*[35]

Démonstration. On a $(h, h') = hh' \cup h'h$. En effet $hh' \cup h'h \supseteq h \cup h'$ et, en vertu de l'infra-invariance de h et de h', $(hh' \cup h'h)(hh' \cup h'h) = hh'hh' \cup hh'hh' \cup h'hhh' \cup h'hh'h = h \cdot h'hh' \cup hh'h \cup h'hh' \cup h' \cdot hh'h = h(h'h \cup hh') \cup hh'h \cup h'hh' \cup h'(hh' \cup h'h) = hh'h \cup hhh' \cup hh'h \cup h'hh' \cup h'hh' \cup h'h'h = hh'h \cup hh' \cup h'hh' \cup h'h = (hh' \cup h'h) \cup hh' \cup (hh' \cup h'h) \cup h'h = hh' \cup h'h$, donc $(h, h') \subseteq hh' \cup h'h$. Et, d'autre part, $hh' \cup h'h \subseteq (h, h')(h, h') \cup (h, h')(h, h') = (h, h') \cup (h, h') = (h, h')$.

Supposons que $c \in H$ est un g-élément pour un des h, h'. Alors il est un g- ou un c-élément pour l'autre. Donc $(hh' \cup h'h)c = hh'c \cup h'hc \supseteq hch' \cup h'ch \supseteq chh' \cup c'h'h = c(hh' \cup h'h)$. Si c est un d-élément pour un des h, h', il est un d- ou un c-élément pour l'autre, et on a de la même manière $c(hh' \cup h'h) \supseteq (hh' \cup h'h)c$.

Enfin si c est un c-élément et pour h et pour h', on a $c(hh' \cup h'h) = (hh' \cup h'h)c$. Donc (h, h') est infra-invariant dans H, et, d'ailleurs, semblable à h et h'. Il est semi-invariant à droite, ou à gauche ou invariant quand ils le sont.

[35] Dans le cas où h et h' sont invariants dans H, ce théorème avait été prouvé par MM. Ore et Dresher (D.O., p. 725); dans un cas particulier, par M. Kuntzmann (Ku., p. 1788).

On voit, de plus, que si (A, B), (A', B') sont les types propres des h, h' sur H, le type propre de (h, h') sur H est un sous-type de $(A \cup A', B \cup B')$.

Il est commode d'introduire pour les sous-hypergroupes infra-invariants reversibles h d'un hypergroupe H une notion nouvelle, celle de *quotient global* H/h *de* H *par* h. Pour cela considérons un élément $c_1 = C_1/h^{(g)}$ de $H/h^{(g)}$ et un élément $c_2 = C_2/h^{(d)}$ de $H/h^{(d)}$. En vertu de l'infra-invariance et de la reversibilité de h, un et un seul des cas suivants doit se présenter: $C_1 \supset C_2$, $C_1 = C_2$, $C_1 \subset C_2$, $C_1 \cap C_2 = 0$. On écrira dans ces cas respectifs $c_1 > c_2$, $c_1 = c_2$, $c_1 < c_2$, $c_1 \| c_2$, et la relation ainsi écrite sera dite la *relation typique* entre c_1 et c_2. L'ensemble T des relations typiques de tous les couples ordonnés (c_1, c_2) $(c_1 \in H/h^{(g)}, c_2 \in H/h^{(d)})$ sera dit *l'ensemble typique* de H pour h. On appellera *quotient global de* H *par* h le complexe $(H/h^{(g)}, H/h^{(d)}, T)$. Deux quotients globaux $H_1/h_1 = (H_1/h_1^{(g)}, H_1/h_1^{(d)}, T_1)$ et $H_2/h_2 = (H_2/h_2^{(g)}, H_2/h_2^{(d)}, T_2)$ sont dits *isomorphes* s'il existe un isomorphisme E_g de $H_1/h_1^{(g)}$ à $H_2/h_2^{(g)}$ et un isomorphisme E_d de $H_1/h_1^{(d)}$ à $H_2/h_2^{(d)}$ tels que, pour tous $c_1 \in H_1/h_1^{(g)}$, $c_2 \in H_1/h_1^{(d)}$, la relation typique entre $E_g c_1$ et $E_d c_2$ soit la même que celle entre c_1 et c_2. Le couple (E_g, E_d) sera dit *un isomorphisme de* H_1/h_1 *à* H_2/h_2.

Manifestement, si (A_1, B_1) et (A_2, B_2) sont les types propres de h_1 sur H_1 et de h_2 sur H_2, on a $A_2/h_2^{(\alpha)} = E_\alpha(A_1/h_1^{(\alpha)})$ et $B_2/h_2^{(\alpha)} = E_\alpha(B_1/h_1^{(\alpha)})$.

Si $H_1/h_1 \simeq H_2/h_2$, manifestement $H_1/h_1^{(g)} \simeq H_2/h_2^{(g)}$, $H_1/h_1^{(d)} \simeq H_2/h_2^{(d)}$ et aussi $H_1/h_1^{(e)} \simeq H_2/h_2^{(e)}$.

THÉORÈME 2. *Soit h un sous-hypergroupe reversible infra-invariant de H, et soit h' un sous-hypergroupe clos de H et non disjoint avec h. Alors $h \cap h'$ est un sous-hypergroupe reversible infra-invariant de h'. hh' est égal à $h'h$ et est le plus petit surhypergroupe commun de h et de h', et on a $hh'/h \simeq h'/(h \cap h')$, l'isomorphisme global précédent pouvant se réaliser en faisant correspondre à une classe C^* suivant $h^{(g)}$ ou $h^{(d)}$ la classe $C^* \cap h'$ suivant $(h \cap h')^{(g)}$ respectivement $(h \cap h')^{(d)}$.*

Un sous-ensemble $\bar{h} \supset h$ de hh' est un sous-hypergroupe de hh' si, et seulement si, $\bar{h} \cap h'$ en est un de h'. $\bar{h} \cap h'$ est reversible et infra-invariant dans h' si $\bar{h}$ l'est dans hh'. Et dans ce cas, la correspondance précédente fournit un isomorphisme global de $\bar{h}/h$ à $(\bar{h} \cap h')/(h \cap h')$. $(\bar{A}^*, \bar{B}^*)$ et $(\bar{A}, \bar{B})$ étant respectivement les types propres de $\bar{h}$ sur hh' et de $h \cap h'$ sur h', on a $\bar{A} = \bar{A}^* \cap h'$, $\bar{B} = \bar{B}^* \cap h'$, $\bar{A}^* = h\bar{A} = \bar{A}h$, $\bar{B}^* = h\bar{B} = \bar{B}h$.[36]

Démonstration. (1) En vertu d'un résultat cité de MM. Ore et Dresher, $h \cap h'$ est un sous-hypergroupe reversible de h'. Soit $c \in h'$, et soit $C = c(h \cap h')$. Posons $C^* = Ch = ch$. Puisque $c \in h'$ et puisque h' est clos, on a $C^* \cap h' = c(h \cap h') = C$. Donc, aux C différents correspondent des C^* différents. De même, si l'on pose $C = (h \cap h')c$ $(c \in h')$ et $C^* = hC = hc$,

[36] Dans le cas où h et $\bar{h}$ sont invariants, ce théorème avait été prouvé par MM. Dresher et Ore (D.O., p. 724 et pp. 726–727). Dans le cas où, de plus, tous les hypergroupes de l'énoncé sont normaux, il avait été prouvé par M. Kuntzmann (Ku., p. 1788). Des formes affaiblies de ce théorème sous d'autres conditions avaient été prouvées par M. Marty (M.2, pp. 96–99).

on trouve de la même manière que $C^* \cap h' = C$. Donc encore, aux C différents correspondent des C^* différents.

Enfin, soit $hch \subseteq hc \cup ch$ $(c \in h')$. Alors $(h \cap h')c(h \cap h') \subseteq hch \cap h' \subseteq (hc \cup ch) \cap h' = (hc \cap h') \cup (ch \cap h') = (h \cap h')c \cup c(h \cap h')$.

Donc si h est infra-invariant dans H, $h \cap h'$ l'est dans h'.

On voit que la correspondance $C^* \to C$ met $h'h/h^{(g)}$, $h'h/h^{(d)}$ en correspondance biunivoque avec respectivement $h'/(h \cap h')^{(g)}$, $h'/(h \cap h')^{(d)}$. Et si $\bar{h}$ est un sous-hypergroupe quelconque de $hh' \cap h'h$, contenant h, $(\bar{h} \cap h')/(h \cap h')^{(\alpha)}$ correspond à $\bar{h}/h^{(\alpha)}$ dans les correspondances biunivoques précédentes pour tout $\alpha = g, d$.

Pour prouver que cette correspondance biunivoque est un isomorphisme, il suffit de prouver que quelque soient les classes C_1, C_2 d'un même côté suivant $h \cap h'$ dans h', on a $C_1^*C_2^* \cap h' \subseteq C_1C_2$.

Regardons les cas de classes droites et de classes gauches:

$$(1) \qquad C_1 = c_1(h \cap h'), \qquad C_2 = c_2(h \cap h') \qquad (c_1, c_2 \in h').$$

Alors $C_1^*C_2^* \cap h' = c_1hc_2h \cap h'$. Or hc_2h est égal ou bien à hc_2, ou bien à c_2h. Donc $c_1hc_2h \cap h'$ est égal ou bien à $c_1hc_2 \cap h' = c_1(hc_2 \cap h') = c_1((h \cap h')c_2) = c_1(h \cap h')c_2 \subseteq C_1C_2$ ou bien à $c_1c_2h \cap h' = c_1c_2(h \cap h') \subseteq C_1C_2$ (parce que $c_1c_2 \subseteq h'$) et, toujours, $C_1^*C_2^* \cap h' \subseteq C_1C_2$.

$$(2) \qquad C_1 = (h \cap h')c_1, \qquad C_2 = (h \cap h')c_2 \qquad (c_1, c_2 \in h').$$

On démontre ce cas par un raisonnement analogue appliqué à hc_1h.

Or, $h'/(h \cap h')^{(\alpha)}$ est un hypergroupe. Donc $hh'/h^{(g)}$, $h'h/h^{(d)}$ le sont aussi. Donc $h'h$, hh' sont des sous-hypergroupes de H. Comme ils contiennent tous h et h' et sont contenus dans tout sous-hypergroupe de H contenant h et h', ils sont égaux, et $hh' = h'h$ est le plus petit surhypergroupe commun (h, h') de h et de h'. On a $(h, h')/h^{(\alpha)} \simeq h'/(h \cap h')^{(\alpha)}$ $(\alpha = g, d)$.

En particulier, on voit que, pour tout sous-hypergroupe $\bar{h} \supseteq h$ de $(h, h') = hh' = h'h$, $(\bar{h} \cap h')/(h \cap h')^{(\alpha)}$ $(\alpha = g, d)$ correspond à $\bar{h}/h^{(\alpha)}$. Mais, α étant un quelconque des g, d, un sous-ensemble $\bar{h} \supseteq h$ de (h, h') est un sous-hypergroupe de H si, et seulement si, $\bar{h}/h^{(\alpha)}$ a un sens et est un sous-hypergroupe de $(h, h')/h^{(\alpha)}$. De même, α étant un quelconque des g, d, $\bar{h} \cap h'$ est un sous-hypergroupe de h' si, et seulement si, $(\bar{h} \cap h')/(h \cap h')^{(\alpha)}$ en est un de $h'/(h \cap h')^{(\alpha)}$. Or, puisque $\bar{h}/h^{(\alpha)} \simeq (\bar{h} \cap h')/(h \cap h')^{(\alpha)}$,[37] ils sont ou ne sont pas des hypergroupes en même temps. Donc $\bar{h} \cap h'$ est un sous-hypergroupe de h' si, et seulement si, $\bar{h}$ en est un de H.

En vertu du lemme 1, $h \cap h'$ est reversible dans h' si, et seulement si, pour $\alpha = g$ et pour $\alpha = d$, $(\bar{h} \cap h')/(h \cap h')^{(\alpha)}$ l'est dans $h'/(h \cap h')^{(\alpha)}$, et, de même, $\bar{h}$ est reversible dans (h, h') si, et seulement si, pour $\alpha = g$ et pour $\alpha = d$, $\bar{h}/h^{(\alpha)}$

[37] Cette formule signifie que si $\bar{h}/h^{(\alpha)}$ a un sens, $(\bar{h} \cap h')/(h \cap h')^{(\alpha)}$ a aussi un sens; et qu'il existe un isomorphisme de $(h, h')/h^{(\alpha)}$ à $h'/(h \cap h')^{(\alpha)}$ qui fait correspondre $(\bar{h} \cap h')/(h \cap h')^{(\alpha)}$ à $\bar{h}/h^{(\alpha)}$.

est reversible dans $(h, h')/h^{(\alpha)}$. Donc, puisque $(h, h')/h^{(\alpha)} \simeq h'/(h \cap h')^{(\alpha)}$ et puisque dans cet isomorphisme $(\bar{h} \cap h')/(h \cap h')^{(\alpha)}$ correspond à $\bar{h}/h^{(\alpha)}$, $\bar{h} \cap h'$ est reversible dans h' si, et seulement si, $\bar{h}$ l'est dans (h, h').

Si $c \in h'$, on a, manifestement, $c\bar{h} \cap h' = c(\bar{h} \cap h')$ et, de même, $\bar{h}c \cap h' = (\bar{h} \cap h')c$. Donc, on a $c(\bar{h} \cap h') \subseteq (\bar{h} \cap h')c$ ou $c(\bar{h} \cap h') \supseteq (\bar{h} \cap h')c$ si $c\bar{h} \subseteq \bar{h}c$ respectivement $c\bar{h} \supseteq \bar{h}c$. De plus si $\bar{h}c$ contient deux classes droites disjointes $c_1\bar{h}$ et $c_2\bar{h}$ $(c_1, c_2 \in h')$ suivant $\bar{h}$ dans (h, h'), $(\bar{h} \cap h')c$ contient deux classes droites disjointes $c_1(\bar{h} \cap h')$ et $c_2(\bar{h} \cap h')$ suivant $\bar{h} \cap h'$ dans h', et, de même, si $c\bar{h}$ contient deux classes gauches disjointes suivant $\bar{h}$, $c(\bar{h} \cap h')$ contient deux classes gauches disjointes suivant $\bar{h} \cap h'$.

Il en résulte que $\bar{h} \cap h'$ est infra-invariant dans h' si $\bar{h}$ l'est dans (h, h'); et que C^* désignant une classe quelconque à gauche ou à droite suivant $\bar{h}$ dans (h, h'), on a la correspondance

$$C^* \rightarrow C \cap h',$$

et dans ce cas un isomorphisme global de $(h, h')/\bar{h}$ à $h'/(\bar{h} \cap h')$. $(\bar{A}^*, \bar{B}^*)$ et $(\bar{A}, \bar{B})$ étant les types propres de $\bar{h}$ dans (h, h') et de $\bar{h} \cap h'$ dans h', on voit que dans cette correspondance $\bar{A}$ correspond à $\bar{A}^*$ et $\bar{B}$ correspond à $\bar{B}^*$. Donc $\bar{A} = \bar{A}^* \cap h'$, $\bar{A}^* = h\bar{A} = \bar{A}h$, et $\bar{B} = \bar{B}^* \cap h'$, $\bar{B}^* = h\bar{B} = \bar{B}h$.

6. **Suites infra-normales. Suites de composition. La loi de raffinement de suites infra-normales et la loi de Jordan-Hölder pour les hypergroupes.** Soient $\bar{H}$ et $\bar{\bar{H}} \subset \bar{H}$ deux sous-hypergroupes d'un hypergroupe H. Une chaîne reversible d'hypergroupes (au sens strict)

$$(S) \qquad \bar{H} = H_0, H_1, H_2, \cdots, H_s = \bar{\bar{H}}$$

s'appelle *suite infra-normale entre* $\bar{H}$ *et* $\bar{\bar{H}}$, si, pour tout $i = 1, 2, \cdots, s$, H_i est infra-invariant dans H_{i-1}. (A_i, B_i) étant le type propre de H_i sur H_{i-1}, $\left(\bigcup_{i=1}^{s} A_i, \bigcup_{i=1}^{s} B_i\right)$ s'appelle *type propre* de la suite (S). Si (A, B) est un surtype dans H du type propre de la suite (S), la suite (S) sera dite infra-normale du type (A, B). En particulier, une suite infra-normale du type $(H, 0)$, $(0, H)$, $(0, 0)$ sera dite suite respectivement *semi-normale à droite, semi-normale à gauche, normale.* La suite

$$H_0/H_1, H_1/H_2, \cdots, H_{s-1}/H_s$$

est dite *la suite de quotients globaux de la suite* (S).

La suite infra-normale (S) sera dite une *suite de composition entre* $\bar{H}$ *et* $\bar{\bar{H}}$ si, pour aucun $i = 1, 2, \cdots, s$, il n'existe d'hypergroupe H^* de H tel que $H_{i-1} \supset H^* \supset H_i$ et que H^* soit reversible, infra-invariant et semblable à H_i sur H_{i-1}. Une suite semi-normale à gauche respectivement à droite entre $\bar{H}$ et $\bar{\bar{H}}$ (S) sera dite une *suite de composition gauche* respectivement *droite* entre $\bar{H}$ et $\bar{\bar{H}}$ s'il n'existe aucune suite semi-normale à gauche respectivement à droite entre $\bar{H}$ et $\bar{\bar{H}}$, autre que (S), dont (S) est une suite partielle. De même, une suite normale entre $\bar{H}$ et $\bar{\bar{H}}(S)$ est dite une *suite de composition bilatère* entre $\bar{H}$ et $\bar{\bar{H}}$

s'il n'existe aucune suite normale entre $\bar{H}$ et $\bar{\bar{H}}$ distincte de (S) et dont (S) est une suite partielle. Deux suites infra-normales entre $\bar{H}$ et $\bar{\bar{H}}$ seront dites *sans intersection* s'il n'y a aucun terme commun sauf $\bar{H}$ et $\bar{\bar{H}}$.

LEMME 6. *Si $\bar{H}$ a un nombre fini d'éléments, une suite infra-normale entre $\bar{H}$ et $\bar{\bar{H}}$ est une suite partielle d'une suite de composition entre $\bar{H}$ et $\bar{\bar{H}}$.*

Démonstration. L'ensemble de toutes les suites infra-normales entre $\bar{H}$ et $\bar{\bar{H}}$ dont la suite donnée est une suite partielle est fini et non vide. Donc, il existe dans cet ensemble une suite

$$(S) \qquad \bar{H} = H_0, H_1, H_2, \cdots, H_s = \bar{\bar{H}}$$

dont la longueur s est maximum. Il est impossible que pour un i, $0 < i \leqq s$, il existe un sous-hypergroupe reversible, infra-invariant H^* de H_{i-1}, tel que $H_{i-1} \supset H^* \supset H_i$, et que H^* soit sur H_{i-1} semblable à H_i, parce que dans ce cas la suite

$$H_0, H_1, \cdots, H_{i-1}, H^*, H_i, \cdots, H_s$$

de longueur $s + 1$ ferait encore partie du même ensemble. Donc (S) est une suite de composition entre $\bar{H}$ et $\bar{\bar{H}}$.

LEMME 7. *Si (S) est une suite infra-normale d'un type (A, B) entre $\bar{H}$ et $\bar{\bar{H}}$, et si H' est un sous-hypergroupe clos de H, la suite formée par les termes inégaux de la suite*

$$H' \cap \bar{H} = H' \cap H_0, H' \cap H_1, H' \cap H_2, \cdots, H' \cap H_s = H' \cap \bar{\bar{H}}$$

est une suite infra-normale du type $(H' \cap A, H' \cap B)$ entre $H' \cap \bar{H}$ et $H' \cap \bar{\bar{H}}$.

Démonstration. Soit $H' \cap H_{i-1} \neq H' \cap H_i$. Puisque H' est clos dans H, $H' \cap H_{i-1}$ est clos dans H_{i-1}. D'autre part, H_i est reversible et infra-invariant dans H_{i-1}. Donc, en vertu du théorème 2, $H' \cap H_i = (H' \cap H_{i-1}) \cap H_i$ est reversible et infra-invariant dans $H' \cap H_{i-1}$, et son type sur $H' \cap H_{i-1}$ coïncide avec celui de H_i et le lemme est démontré.

LEMME 8. *Si (S) est une suite infra-normale entre $\bar{H}$ et $\bar{\bar{H}}$ et si H^* est un sous-hypergroupe infra-invariant reversible de $\bar{H}$ dont le type propre est compatible avec celui de la suite (S), pour tout $i = 0, 1, \cdots, s$, $H_iH^* = H^*H_i$ et est un sous-hypergroupe de H, et la suite formée par tous les termes inégaux de la suite*

$$(S^*) \qquad \bar{H} = H_0H^*, H_1H^*, \cdots, H_sH^* = \bar{\bar{H}}H^*$$

est une suite infra-normale entre $\bar{H}$ et $\bar{\bar{H}}H^*$ dont le type propre est compatible avec celui de H^*. Si la suite (S) et H^* sont du type $(\bar{H}, 0)$, $(0, \bar{H})$, $(0, 0)$ sur $\bar{H}$, (S^*) est aussi du même type.

Démonstration. Pour tout $i = 1, 2, \cdots, s$, H_i est clos dans H_{i-1}. Donc, en vertu du lemme 1, H_i est clos dans $H_0 = \bar{H}$. Donc, en vertu du théorème 2, $H_iH^* = H^*H_i$ et est un sous-hypergroupe de $\bar{H}$. $H_{i-1} \cap H^*$ est un sous-hypergroupe reversible infra-invariant de H_{i-1} et, (A^*, B^*) désignant le type propre de H^* sur $\bar{H}$, $H_{i-1} \cap H^*$ est du type (A^*, B^*) sur H_{i-1}. Comme le

type de H_i sur H_{i-1} y est compatible avec (A^*, B^*), $H_iH^* \cap H_{i-1} = H_i(H^* \cap H_{i-1})$ est, en vertu du théorème 2, un sous-hypergroupe reversible et infra-invariant de H_{i-1} d'un type (A_i, B_i) compatible avec (A^*, B^*) sur H_{i-1}. Donc, en vertu du théorème 2, H_iH^* est reversible et infra-invariant dans $H_{i-1}H^*$, et son type propre (A_i^*, B_i^*) sur $H_{i-1}H^*$ est tel que $A_i^*H^* = A_i^*$, $A_i^* \cap H_{i-1} = A_i$ et $B_i^*H^* = B_i^*$, $B_i^* \cap H_{i-1} = B_i$. Comme $A^*H^* = A^*$, $B^*H^* = B^*$, on a $A_i^* \cap B^* = (A_i^* \cap B^* \cap H_{i-1})H^* = (A_i \cap B^*)H^*$ et $B_i^* \cap A^* = (B_i \cap A^*)H^*$. $A_i \cap B^*$ et $A^* \cap B_i$ étant vides, $A_i^* \cap B^*$ et $A^* \cap B_i^*$ le sont aussi, et (A_i^*, B_i^*), donc aussi $(\bigcup_i A_i^*, \bigcup_i B_i^*)$ est compatible avec (A^*, B^*), ce qui prouve le lemme.

En particulier, si (S) est une suite semi-normale d'un certain côté respectivement normale entre $\bar{H}$ et $\bar{\bar{H}}$, et si H^* est semi-invariant dans $\bar{H}$ du même côté respectivement invariant dans $\bar{H}$, $H_{i-1} \cap H^*$ et $H_i(H_{i-1} \cap H^*) = H_iH^* \cap H_{i-1}$ est semi-invariant du même côté respectivement invariant dans H_{i-1}, et H_iH^* l'est dans $H_{i-1}H^*$, donc la suite (S^*) est semi-normale respectivement normale.

Soient

$$(S) \qquad \bar{H} = H_0, H_1, H_2, \cdots, H_s = \bar{\bar{H}},$$

$$(S') \qquad \bar{H} = H'_0, H'_1, H'_2, \cdots, H'_{s'} = \bar{\bar{H}}$$

deux suites infra-normales entre $\bar{H}$ et $\bar{\bar{H}}$ dont les types sont compatibles. La suite des hypergroupes distincts parmi les $\mathcal{H}_{i,j} = (H_{i-1} \cap H'_j)H_i = H_i(H_{i-1} \cap H'_j)$ rangés suivant l'ordre lexicographique de leur indices, sera appelée *le raffinement de la suite* (S) *par la suite* (S'). Le raffinement de (S) par (S') est formé par la juxtaposition des suites Σ_i, Σ_i étant la suite des hypergroupes distincts parmi les $\mathcal{H}_{i,j}$ $(j = 1, 2, \cdots, s')$.

En vertu du lemme 7 la suite des hypergroupes distincts parmi les $H_{i-1} \cap H'_j$ est une suite infra-normale entre H_{i-1} et H dont le type est compatible avec celui de la suite (S) et, en particulier, avec celui de H_i sur H_{i-1}. Donc Σ_i est, en vertu du lemme 8, une suite infra-normale entre H_{i-1} et $\bar{\bar{H}}H_i = H_i$ dont le type est compatible avec celui de H_i sur H_{i-1}. Donc le raffinement de (S) par (S') est une suite infra-normale entre $\bar{H}$ et $\bar{\bar{H}}$, comprenant (S) comme suite partielle, et dont le type est compatible avec celui de (S). En particulier, si (S) et (S') sont des suites semi-normales d'un même côté respectivement normales, le raffinement de (S) par (S'), comme on voit facilement, est une suite semi-normale du même côté respectivement normale.

(S) et (S') sont dites *sans raffinement mutuel* si le raffinement de (S) par (S') coïncide avec (S), et si le raffinement de (S') par (S) coïncide avec (S'). Manifestement, si (S) et (S') sont deux suites de composition (dont les types sont compatibles) ou de composition gauche, ou de composition droite, ou de composition bilatère, elles sont sans raffinement mutuel.

LEMME 9. $(H_{i-1} \cap H'_j)H_i \cap (H_{i-1} \cap H'_{j-1}) = (H_i \cap H'_{j-1})H'_j \cap (H_{i-1} \cap H'_{j-1}) = (H_{i-1} \cap H'_j)(H_i \cap H'_{j-1})$ *quand* (S) *et* (S') *sont deux suites infra-normales entre* $\bar{H}$ *et* $\bar{\bar{H}}$ *dont les types sont compatibles.*

Démonstration. En effet, dans ce cas $(H_{i-1} \cap H'_j)H_i$ et $H_{i-1} \cap H'_{j-1}$ sont clos dans H_{i-1}, donc puisque $H_{i-1} \cap H'_j \subseteq H_{i-1} \cap H'_{j-1}$, on a

$$\begin{aligned}(H_{i-1} \cap H'_j)H_i \cap (H_{i-1} \cap H'_{j-1}) &= (H_{i-1} \cap H'_j)(H_i \cap (H_{i-1} \cap H'_{j-1})) \\ &= (H_{i-1} \cap H'_j)(H_i \cap H'_{j-1}),\end{aligned}$$

et, de même, $H'_j(H_i \cap H'_{j-1}) \cap (H_{i-1} \cap H'_{j-1}) = (H_{i-1} \cap H'_j)(H_i \cap H'_{j-1})$.

THÉORÈME 3 (Loi de raffinement pour les suites infra-normales).[38] (S) *et* (S') *étant deux suites infra-normales entre* $\bar{H}$ *et* $\bar{\bar{H}}$ *dont les types sont compatibles, le raffinement* (Σ) *de* (S) *par* (S') *et le raffinement* (Σ') *de* (S') *par* (S) *ont la même longueur et, à l'ordre et à l'isomorphisme global près, la même suite de quotients globaux. De plus* $(\mathfrak{A}, \mathfrak{B})$ *et* $(\mathfrak{A}', \mathfrak{B}')$ *étant les types propres de* (Σ) *et de* (Σ'), *on a*

$$\mathfrak{A}' = \bigcup_{i=1}^{s} \bigcup_{j=1}^{s'} [(\mathfrak{A} \cap [(H_{i-1} \cap H'_{j-1}) - (H_{i-1} \cap H'_j)(H_i \cap H'_{j-1})])H'_j],$$

$$\mathfrak{B}' = \bigcup_{i=1}^{s} \bigcup_{j=1}^{s'} [(\mathfrak{B} \cap [(H_{i-1} \cap H'_{j-1}) - (H_{i-1} \cap H'_j)(H_i \cap H'_{j-1})])H'_j].$$

Démonstration. Puisque $(H_{i-1} \cap H'_j)H_i$ est infra-invariant dans $(H_{i-1} \cap H'_{j-1})H_i$ $(0 < i \leqq s, 0 < j \leqq s')$, et puisque $H_{i-1} \cap H'_{j-1}$ est clos dans $\bar{H}$, on a, en vertu du théorème 2,

$$\begin{aligned}&(H_{i-1} \cap H'_{j-1})H_i/(H_{i-1} \cap H'_j)H_i \\ &\qquad = (H_{i-1} \cap H'_{j-1})[(H_{i-1} \cap H'_j)H_i]/(H_{i-1} \cap H'_j)H_i \\ &\qquad \simeq (H_{i-1} \cap H'_{j-1})/[(H_{i-1} \cap H'_j)H_i \cap (H_{i-1} \cap H'_{j-1})] \\ &\qquad = (H_{i-1} \cap H'_{j-1})/(H_{i-1} \cap H'_j)(H_i \cap H'_{j-1}),\end{aligned}$$

et en vertu du même théorème, le type propre de $(H_{i-1} \cap H'_j)(H_i \cap H'_{j-1})$ sur $H_{i-1} \cap H'_{j-1}$ est

$$\begin{aligned}(\bar{\mathfrak{A}}, \bar{\mathfrak{B}}) = (&\mathfrak{A} \cap [(H_{i-1} \cap H'_{j-1}) - (H_{i-1} \cap H'_j)(H_i \cap H'_{j-1})], \\ &\mathfrak{B} \cap [(H_{i-1} \cap H'_{j-1}) - (H_{i-1} \cap H'_j)(H_i \cap H'_{j-1})]).\end{aligned}$$

On a $\bar{\mathfrak{A}}(H_i \cap H'_{j-1}) = \bar{\mathfrak{A}}$, $\bar{\mathfrak{B}}(H_i \cap H'_{j-1}) = \bar{\mathfrak{B}}$. Mais, en vertu du même théorème, puisque $(H_i \cap H'_{j-1})H'_j$ est infra-invariant dans $(H_{i-1} \cap H'_{j-1})H'_j$, on a

[38] La démonstration de ce théorème et des lemmes précédents est à peu près semblable à la démonstration de la loi de raffinement pour les groupes par la méthode de Zassenhaus. Les seules complications proviennent de ce que les chaînes doivent être closes et de ce que certains types doivent être compatibles pour permettre l'application du théorème 2.

Le présent théorème avait été prouvé par MM. Ore et Dresher pour le cas des suites normales (voir D.O., p. 727); auparavant, il avait été prouvé par M. Kuntzmann pour les suites de composition bilatères dont tous les termes sont des hypergroupes normaux au sens de M. Marty (voir Ku., p. 1788).

$$
\begin{aligned}
(H_{i-1} \cap H'_{j-1})/(H_{i-1} \cap H'_j)(H_i \cap H'_{j-1}) \\
&= (H_{i-1} \cap H'_{j-1})/[(H_i \cap H'_{j-1})H'_j \cap (H_{i-1} \cap H'_{j-1})] \\
&\simeq (H_{i-1} \cap H'_{j-1})[(H_i \cap H'_{j-1})H'_j]/(H_i \cap H'_{j-1})H'_j \\
&= (H_{i-1} \cap H'_{j-1})H'_j/(H_i \cap H'_{j-1})H'_j\,,
\end{aligned}
$$

et le type de $(H_i \cap H'_{j-1})H'_j$ sur $(H_{i-1} \cap H'_{j-1})H'_j$ est, en vertu du même théorème, $(\bar{\mathfrak{A}}(H_i \cap H'_{j-1})H'_j\,, \mathfrak{B}(H_i \cap H'_{j-1})H'_j) = (\bar{\mathfrak{A}}H'_j\,, \mathfrak{B}H'_j) = ((\mathfrak{A} \cap [(H_{i-1} \cap H'_{j-1}) - (H_{i-1} \cap H'_j)(H_i \cap H'_{j-1})])H'_j\,, (\mathfrak{B} \cap [(H_{i-1} \cap H'_{j-1}) - (H_{i-1} \cap H'_j)(H_i \cap H'_{j-1})])H'_j)$. Par conséquent $(H_{i-1} \cap H'_{j-1})H_i/(H_{i-1} \cap H'_j)H_i \simeq (H_{i-1} \cap H'_{j-1})H'_j/(H_i \cap H'_{j-1})H'_j$ et $(H_{i-1} \cap H'_{j-1})H'_j = (H_i \cap H'_{j-1})H'_j$ si, et seulement si, $(H_{i-1} \cap H'_{j-1})H_i = (H_{i-1} \cap H'_j)H_i$, ce qui montre que les suites des quotients globaux de (Σ) et de (Σ') sont les mêmes à l'ordre et à l'isomorphisme global près. Et le type propre $(\mathfrak{A}', \mathfrak{B}')$ de (Σ') est bien donné par la formule de l'énoncé.

Théorème 4 (Loi de Jordan-Hölder pour les hypergroupes). *Si (S) et (S') sont deux suites infra-normales à types propres compatibles sans raffinement mutuel (en particulier deux suites de composition, ou de composition droite, ou de composition gauche, ou de composition bilatère) entre $\bar{H}$ et $\bar{\bar{H}}$, elles ont la même longueur et, à l'ordre et à l'isomorphisme global près, la même suite de quotients globaux. Le type propre de l'une de ces suites se calcule à partir du celui de l'autre par les formules du théorème précédent.*

Démonstration. On n'a qu'appliquer le théorème précédent et remarquer que $(\Sigma) = (S)$, $(\Sigma') = (S')$.

Conséquence. *(S) et (S') étant deux suites infra-normales à types propres compatibles sans raffinement mutuel, si (S) est semi-normale d'un certain côté ou normale, (S') l'est aussi.*

Démonstration. En effet, (A, B) et (A', B') étant les types propres de (S) et (S'), donc aussi de leurs raffinements $(\Sigma) = (S)$ et $(\Sigma') = (S')$, les formules du théorème 3 montrent que si $A = 0$, on a $A' = 0$, et que si $B = 0$, on a $B' = 0$, ce qui prouve la proposition.

Conséquence. *S'il existe une suite de composition (S) entre $\bar{H}$ et $\bar{\bar{H}}$ qui est normale, toutes les suites de composition entre $\bar{H}$ et $\bar{\bar{H}}$ le sont.*

Démonstration. (S') étant une suite de composition quelconque entre $\bar{H}$ et $\bar{\bar{H}}$, le type propre de (S') est certainement compatible avec le type propre $(0, 0)$ de (S) et, de plus, (S) et (S') sont sans raffinement mutuel. Donc (S') est normale.

Cette conséquence subsiste encore quand on y remplace "de composition" par "de composition gauche" ou "de composition droite".

Soient (S) et (S') deux suites infra-normales entre $\bar{H}$ et $\bar{\bar{H}}$ types propres compatibles et sans raffinement mutuel. s étant leur longueur commune et i

étant un des nombres 1, 2, $\cdots$, s, soit $\varphi(i)$ le plus petit nombre j, $0 < j \leqq s$, tel que $H'_j \cap H_{i-1} = H'_j \cap H_i$. En vertu du lemme 9, i est aussi le plus petit indice tel que $H_i \cap H'_{\varphi(i)-1} = H_i \cap H'_{\varphi(i)}$. On dira que les indices i_1, i_2 sont *en inversion* (pour le couple ordonné $[(S), (S')]$), si $(i_1 - i_2)\,[\varphi(i_1) - \varphi(i_2)] < 0$. Si les suites (S) et (S') ne coïncident pas, il y a des couples d'indices en inversion: parce que s'il n'y en avait pas, on aurait $\varphi(i) = i$. Donc si i était le plus petit indice tel que $H_i \neq H'_i$ (donc $H_{i-1} = H'_{i-1}$) on aurait $H_i = H_i \cap H_{i-1} = H_i \cap H'_{i-1} = H_i \cap H'_i = H_{i-1} \cap H'_i = H'_i$, ce qui est absurde.

Lemme 10. *Si les indices i, $i_1 > i$ sont en inversion et si $j = \varphi(i_1)$, on a*

$$H_{i-1} \cap H'_{j-1} = (H_{i-1} \cap H'_j)(H_i \cap H'_{j-1}),$$

$$(H_{i-1} \cap H'_j)H_i = H_{i-1}, \qquad (H_i \cap H'_{j-1})H'_j = H'_{j-1}.$$

Démonstration. Puisque $i < i_1$, et puisque i et i_1 sont en inversion, on a $j = \varphi(i_1) < \varphi(i)$, donc $H_{i-1} \cap H'_j \neq H_i \cap H'_j$ et $(H_{i-1} \cap H'_j)H_i \neq H_i$. Puisque (S') ne raffine pas (S), on doit avoir $(H_{i-1} \cap H'_j)H_i = H_{i-1}$ et, de même $(H_i \cap H'_{j-1})H'_j = H'_{j-1}$. Donc $(H_{i-1} \cap H'_j)(H_i \cap H'_{j-1}) = (H_{i-1} \cap H'_j)H_i \cap H'_{j-1} = H_{i-1} \cap H'_{j-1}$, et tout est prouvé.

Théorème 5. *Si i et i_1 sont en inversion, ou bien H_i est invariant dans H_{i-1}, ou bien H_{i_1} est invariant dans H_{i_1-1}, ou bien H_i dans H_{i-1} et H_{i_1} dans H_{i_1-1} sont semi-invariants d'un même côté.*

Démonstration. Puisque $(H_{i-1} \cap H'_{j-1})H_i = H_{i-1}$, H_i est invariant respectivement semi-invariant d'un certain côté dans H_{i-1} si, et seulement si, $H_i \cap (H_{i-1} \cap H'_{j-1}) = H_i \cap H'_{j-1}$ l'est dans $H_{i-1} \cap H'_{j-1}$. De même H'_j est invariant respectivement semi-invariant d'un certain côté dans H'_{j-1} si, et seulement si, $H_{i-1} \cap H'_j$ l'est dans $H_{i-1} \cap H'_{j-1}$. Enfin, en vertu de la loi Jordan-Hölder $H_{i_1-1}/H_{i_1} \simeq H_{\varphi(i_1)-1}/H_{\varphi(i_1)} = H'_{j-1}/H'_j$, donc H_{i_1} est invariant ou semi-invariant d'un certain côté dans H_{i_1-1} si, et seulement si, H'_j l'est dans H'_{j-1}.

Or, les types propres de H_i dans H_{i-1} et de H'_j dans H'_{j-1} sont compatibles. Puisque, en vertu du théorème 2, ils coïncident sur $H_{i-1} \cap H'_{j-1}$ avec ceux respectivement de $H_i \cap H'_{j-1}$ et de $H_{i-1} \cap H'_j$, ces derniers types sont compatibles. En vertu du lemme précédent, $(H_{i-1} \cap H'_j)(H_i \cap H'_{j-1}) = H_{i-1} \cap H'_{j-1}$. Par conséquent, il résulte du lemme 5 que ou bien $H_{i-1} \cap H'_j$ est invariant dans $H_{i-1} \cap H'_{j-1}$, ou bien $H_i \cap H'_{j-1}$ l'est, ou bien $H_{i-1} \cap H'_j$ et $H_i \cap H'_{j-1}$ sont semi-invariants d'un même côté dans $H_{i-1} \cap H'_{j-1}$, ce qui prouve la proposition.

Conséquence. *Si (S) est une suite infra-normale entre $\bar{H}$ et $\bar{\bar{H}}$, et s'il existe une suite infra-normale entre $\bar{H}$ et $\bar{\bar{H}}$ autre que (S) et dont le type propre est compatible avec celui de (S), ou bien il existe un indice i tel que H_i soit invariant dans H_{i-1}, ou bien il existe deux indices i_1, i_2 tels que H_{i_1} dans H_{i_1-1} et H_{i_2} dans H_{i_2-1} soient semi-invariants d'un même côté.*

En effet, si une telle suite (S') existait, sans que (S) satisfasse à nos conditions, il existerait deux indices i_1, i_2 en inversion pour $((S), (S'))$, et le théorème 5 serait en défaut pour ces indices.

Quand H est un hypergroupe$_G$, ou un hypergroupe$_C$, ou un hypergroupe$_D$, toute chaîne de sous-hypergroupes de H est reversible, donc on peut supprimer la condition de la reversibilité de la châine (S) dans la définition des suites infra-normales entre deux sous-hypergroupes d'un tel hypergroupe. De plus, puisque, dans ce cas, $\{1_H\}$ est infra-invariant dans H, donc aussi dans $\bar{H}$, toute suite infra-normale entre $\bar{H}$ et $\bar{\bar{H}}$ est une suite partielle d'une suite infra-normale entre $\bar{H}$ et $\{1_H\}$. Si $\bar{H}$ a un nombre fini d'éléments, elle est une suite partielle d'une suite de composition entre $\bar{H}$ et $\{1_{\bar{H}}\}$. Si H a un nombre fini d'éléments, les suites de composition entre H et $\{1_H\}$ seront appelées *suites de composition de H*.

Quand H est un hypergroupe$_G$ fini, toutes les suites infra-normales dans H sont semi-normales à gauche, et toutes les suites semi-normales à droite sont normales. De même, quand H est un hypergroupe$_D$ fini, toutes les suites infra-normales dans H sont semi-normales à droite, et toutes les suites semi-normales à gauche sont normales. Donc dans ces deux cas, les types propres de deux suites infra-normales quelconques entre $\bar{H}$ et $\bar{\bar{H}}$ ($H \supseteq \bar{H} \supseteq \bar{\bar{H}}$) sont compatibles, ainsi que les types propres d'une suite infra-normale dans H et d'un sous-hypergroupe infra-invariant de H. Donc, pour des tels hypergroupes, on peut supprimer dans les énoncés du lemme 8 et des théorèmes 3 et 4 toutes les hypothèses concernant les types des suites infra-normales ou des hypergroupes dont il y est question.

De plus dans ce cas, *les suites d'indices des raffinements de deux suites infra-normales entre $\bar{H}$ et H l'une par l'autre sont les mêmes à l'ordre près.* En particulier, *si deux suites infra-normales entre $\bar{H}$ et H sont sans raffinement mutuel* (ce qui a lieu, en particulier, pour deux suites de composition entre $\bar{H}$ et $\bar{\bar{H}}$), *elles ont la même suite d'indices à l'ordre près.*

7. **Métagroupes.** *Un sous-hypergroupe reversible infra-invariant h de H est dit infra-invariant strict dans H si $H/h^{(c)}$ est un groupe.* Soit h un tel sous-hypergroupe de H, et soient c, c' deux éléments de H tels que cc' ne soit pas disjoint avec h, alors $hchc'h = hch \cdot hc'h \supseteq cc'$ n'est pas disjoint avec h. Donc, puisque $h/h^{(c)}$ est l'unité du groupe $H/h^{(c)}$, on doit avoir $hchc'h = h$, c'est-à-dire $chc' \subseteq h$. Donc $hc'h/h^{(c)}$ est l'inverse de $hch/h^{(c)}$ dans $H/h^{(c)}$. Les inverses dans un groupe étant permutables, on doit avoir $hc'h \cdot hch = hc'hch = h$, c'est-à-dire $c'hc \subseteq h$. En particulier $c'c \subseteq h$. c étant un élément quelconque de H, il existe un $c' \in H$ tel que $cc' \cap h$ ne soit pas vide et un $c'' \in H$ tel que $c''c \cap h$ ne soit pas vide. Donc $chc'' \subseteq h$. Donc $chc''c \subseteq hc$. Comme $c''c \subseteq h$, on a $hc''c = h$, et $ch \subseteq hc$. Donc $hc = ch$, et h est *invariant* dans H. *Donc tout sous-hypergroupe infra-invariant strict de H y est invariant.*[39]

Remarque. Si $\bar{\bar{H}}$ est un sous-hypergroupe reversible infra-invariant strict de

[39] MM. Dresher et Ore appellent un sous-hypergroupe invariant h d'un hypergroupe H possèdant des unités à droite et des unités à gauches *rigoureusement invariant* (*strongly normal*) *dans H* si $H/h^{(d)}$ ($= H/h^{(c)}$) est un groupe (voir D.O., pp. 728–731). On voit que pour de tels hypergroupes H les notions de l'invariance stricte au sens de ce travail et de l'invariance rigoureuse au sens de MM. Dresher et Ore coïncident.

$\bar{H}$, et si H', $\bar{H} \supset H' \supset \bar{\bar{H}}$, est un sous-hypergroupe de $\bar{H}$, H' est reversible dans $\bar{H}$. En effet, les $H'/\bar{\bar{H}}^{(\alpha)}$ ($\alpha = g, d$) sont des sous-groupes des $\bar{H}/\bar{\bar{H}}^{(\alpha)}$ correspondants. Ils sont donc reversibles dans les $\bar{H}/\bar{\bar{H}}^{(\alpha)}$. Donc en vertu du lemme 3, H' est reversible dans $\bar{\bar{H}}$.

THÉORÈME 6. *Si h est un sous-hypergroupe d'un hypergroupe H tel que, pour tout $c \in H$,*

(1) *il existe un $\bar{c} \in H$ tel que $c \in h\bar{c}$ et $\bar{c} \in h\bar{c}$;*

(2) *il existe un $c' \in H$ et un $c'' \in H$ tels que $chc' \subseteq h$ et $c''hc \subseteq h$,*

h est reversible et infra-invariant strict dans H.[40]

Démonstration. Soit $c \in H$, et soient c', c'' éléments de H tels que $chc' \subseteq h$, $c''hc \subseteq h$.

On a $hchc' = c''hch = h$. Donc, $c''h = c''hchc' = hc'$. Mais alors $hc = chc'hc = cc''hhc = cc''hc \subseteq ch$ et $ch = chc''hc = chhc'c = chc'c \subseteq hc$. Donc $ch = hc = hch$.

On a $\bar{c} \in h\bar{c} = \bar{c}h$, $hc = ch \subseteq h\bar{c} = \bar{c}h$. Alors, si $\bar{c}'$, $\bar{c}'' \in H$ sont tels que $\bar{c}h\bar{c}' \subseteq h$ et $\bar{c}''h\bar{c} \subseteq h$, on a $ch\bar{c}' \subseteq \bar{c}h\bar{c}' \subseteq h$ et $\bar{c}''hc \subseteq h$. Donc $h\bar{c} = \bar{c}h = \bar{c}h\bar{c}''hc = \bar{c}hh\bar{c}'c = \bar{c}h\bar{c}'c \subseteq hc = ch$. Donc $c \in ch = hc$ et, si $hc \subseteq h\bar{c}$, on a $hc = h\bar{c}$.

Soient c_1, $c_2 \in H$. Supposons que $hc_1 \cap hc_2$ n'est pas vide, et soit $c \in hc_1 \cap hc_2$. On a $hc \subseteq h \cdot hc_1 = hc_1$, donc $hc = hc_1$. De même $hc \subseteq hc_2$, donc $hc = hc_2$. Donc, on voit que $hc_1 = hc_2$, et h est reversible dans H.

Soient c_1, c_2 deux éléments quelconques de H. Soit c_1' un élément de H tel que $c_1hc_1' \subseteq h$. Et soit c^* un élément de H tel que $c_1'c^* \ni c_2$. On a $hc_1h \cdot hc_2h \subseteq hc_1h \cdot hc_1'h \cdot hc^*h = h \cdot hc^*h = hc^*h$. Donc le composé de deux éléments quelconques de $H/h^{(c)}$ est l'ensemble d'un seul élément. $H/h^{(c)}$ est, donc, groupe.

Remarque. La condition (1) de ce théorème est certainement vérifiée si H a un nombre fini d'éléments. En effet formons pour un $c \in H$ une suite d'éléments $c = c_0, c_1, c_2, \cdots, c_i, \cdots$ par le procédé suivant. Supposons qu'on ait déjà formé c_i et supposons qu'il n'est pas un élément de hc_i. Puisque $hH = H$, il existe des éléments c' de H tels que $hc' \ni c_i$. Un de ces c' sera pris pour c_{i+1}. Puisque H a un nombre fini d'éléments on doit arriver à un $\bar{c} = c_m$ tel que $h\bar{c} \ni \bar{c}$. Or, si $c \in hc_i$, on a $c \in hc_i \subseteq h \cdot hc_{i+1} = hc_{i+1}$. Comme $c = c_0 \in hc_1$, on a $c \in hc_m = h\bar{c}$, et $\bar{c}$ satisfait à la condition (1).[41]

On appelle une suite

$$(S) \qquad \bar{H} = H_0 \supset H_1 \supset H_2 \supset \cdots \supset H_s = \bar{\bar{H}}$$

[40] MM. Ore et Dresher avaient prouvé ce théorème dans le cas où H a des unités droites et gauches, mais en remplaçant l'hypothèse (1) par l'hypothèse plus forte que h contient toutes les unités droites et gauches de H et en supposant, de plus, sur les éléments c', c'' de la condition (2) que $c'c$ contient une unité gauche de H et que cc'' contient une unité droite de H. Voir D.O., pp. 728–730.

[41] Il est à remarquer que la condition (1) est verifiée si h (pouvant être infini) contient une unité (droite centrale ou gauche) de H (ce qui n'est, d'ailleurs, possible que si H a de telles unités). En effet, on a vu que la condition (2) seule permet de démontrer que, pour tout $c \in H$, on a $hc = ch$. Si $e \in h$ est une unité gauche (resp. droite) de H, on a $c \in ec \subseteq hc$ (resp. $c \in ce \subseteq ch = hc$). Si e est une unité centrale de H, on a $c \in ece \subseteq hch = hc$.

suite stricte entre $\bar{H}$ et $\bar{\bar{H}}$ si, pour tout $i = 1, 2, \cdots, s$, H_i est un sous-hypergroupe infra-invariant strict de H_{i-1}.

En particulier, s'il n'existe aucune suite stricte entre $\bar{H}$ et $\bar{\bar{H}}$ contenant (S) comme suite partielle et autre que (S), (S) est dite une *suite de composition stricte entre* $\bar{H}$ *et* $\bar{\bar{H}}$.

Si $\bar{H}$ a un nombre fini d'éléments, toute suite de composition stricte entre $\bar{H}$ et $\bar{\bar{H}}$ est une suite de composition entre $\bar{H}$ et $\bar{\bar{H}}$. En effet, si $H_{i-1}/H_i^{(c)}$ est un groupe d'ordre fini, et s'il existe un sous-hypergroupe H', $H_{i-1} \supset H' \supset H_i$, de H_{i-1} infra-invariant dans H_{i-1}, $H'/H_i^{(c)}$ est un sous-hypergroupe infra-invariant du groupe $H_{i-1}/H_i^{(c)}$ c'est-a-dire un sous-groupe invariant de ce groupe. Donc $H_{i-1}/H'^{(c)} \simeq (H_{i-1}/H_i^{(c)})/(H'/H_i^{(c)})^{(c)}$ est un groupe, et H' est strict dans H_{i-1}. D'autre part, si $\bar{\bar{H}}$ est infra-invariant strict et d'indice fini dans $\bar{H}$, il existe une suite de composition stricte entre $\bar{H}$ et $\bar{\bar{H}}$. En effet, le raisonnement précédent montre que, si H', $\bar{H} \supset H' \supset \bar{\bar{H}}$ est un sous-hypergroupe infra-invariant de $\bar{H}$, $\bar{H}/H'^{(c)}$ et $H'/\bar{\bar{H}}^{(c)}$ sont des groupes. Donc, si $(\bar{H}:\bar{\bar{H}})$ est fini on peut intercaler entre $\bar{H}$ et $\bar{\bar{H}}$ des sous-hypergroupes de $\bar{H}$ de manière à former une suite de composition stricte.

THÉORÈME 7. *S'il existe une suite de composition infra-invariante stricte entre* $\bar{H}$ *et* $\bar{\bar{H}}$, *toutes les suites de composition infra-invariantes entre* $\bar{H}$ *et* $\bar{\bar{H}}$ *sont strictes; elles ont toutes, dans ce cas, la même longueur et le même ensemble de quotients.*

Démonstration. La suite de composition infra-invariante stricte entre $\bar{H}$ et $\bar{\bar{H}}$ dont l'existence est supposée est, en particulier, une suite de composition invariante.

Donc toutes les suites de composition entre $\bar{H}$ et $\bar{\bar{H}}$ sont invariantes et ont le même ensemble de quotients formé de groupes. Donc elles sont suites de composition stricte.

Un hypergroupe H s'appelle métagroupe s'il existe un $e \in H$ et une suite stricte

$$(S) \qquad H = H_0 \supset H_1 \supset H_2 \supset \cdots \supset H_s = \{e\}$$

entre H et $\{e\}$. Dans ce cas, si H est d'ordre fini il existe une suite de composition entre H et $\{e\}$ dont l'ensemble de quotients est formé de groupes seuls et ceci est aussi suffisant pour que H soit un métagroupe. Et toutes les suites de composition entre H et $\{e\}$ sont suites de composition stricte.

MM. Ore et Dresher appellent métagroupes possédant une unité *ultragroupes.*[42, 43]

$\bar{\bar{H}}$ étant un sous-hypergroupe reversible de $\bar{H}$, s'il existe une suite stricte entre $\bar{H}$ et $\bar{\bar{H}}$, $\bar{H}/\bar{\bar{H}}^{(\alpha)}$ $(\alpha = c, g, d)$ est un métagroupe et inversement, parce que si

[42] D.O., pp. 732–733. Je trouve, d'ailleurs, que ce nom n'est pas bien choici; en effet, étymologiquement, le préfixe *ultra* placé devant une notion signifié qu'il s'agit d'un cas particulier de cette notion où ses propriétés se réalisent d'une manière particulièrement forte. Ici, il s'agit, au contraire, d'un affaiblissement des propriétés de groupe.

[43] L'élément $e \in H$ est une unité du groupe H_{s-1}, mais non obligatoirement de H. Donc, un métagroupe peut ne pas être un ultragroupe au sens des MM. Dresher et Ore.

$\bar{H} \supset H' \supset H'' \supset \bar{\bar{H}}$, et si H'' est reversible dans H', on a $(H'/\bar{\bar{H}}^{(\alpha)})/(H''/\bar{\bar{H}}^{(\alpha)})^{(\alpha)} \simeq H'/H''^{(\alpha)}$, donc c'est un groupe si $H'/H''^{(\alpha)}$ l'est. Si H est un métagroupe, toute suite stricte entre H et $\bar{H} \ni e$ est une suite partielle d'une suite stricte entre H et $\{e\}$. Et tout sous-hypergroupe clos $\bar{H} \ni e$ de H est un métagroupe. Si H est un hypergroupe$_D$, pour que $h \subset H$ y soit strict, il faut qu'il y soit invariant. Mais, comme j'avais montré, *s'il est invariant dans H, il y est strict.* Donc, $H/h^{(d)}$ *est dans ce cas un métagroupe* (je l'appelle dans ce cas métagroupe$_D$) *si, et seulement s'il existe une suite de composition invariante entre H et h.* On a des résultats analogues pour les hypergroupes$_G$.

L'application de la théorie précédente à la théorie des corps de nombres $\mathfrak{P}$-adiques montre, en même temps, qu'il existe dans les hypergroupes$_D$ de très nombreux sous-hypergroupes semi-invariants (à droite) qui ne sont pas invariants et qui y jouent un rôle important. Corrélativement, on peut conclure qu'il existe dans les hypergroupes$_G$ des sous-hypergroupes semi-invariants à gauche, mais non invariants, importants.

Ceci montre, d'ailleurs, que l'étude précédente a un intérêt qui n'est pas seulement formel.

D'autre part, comme je l'ai déjà mentionné, j'ai trouvé un exemple de sous-groupe infra-invariant, mais non invariant, d'un groupe. Et, un tel sous-groupe ne peut pas être semi-invariant.

PARIS, FRANCE

INVARIANTS OF A SYSTEM OF LINEAR HOMOGENEOUS DIFFERENTIAL EQUATIONS OF THE SECOND ORDER

By I. A. Barnett and H. Reingold

Introduction. E. J. Wilczynski, in his treatise *Projective Differential Geometry of Curves and Ruled Surfaces* (Leipzig, Teubner, 1906), Chapter IV, discusses the invariants of a system of linear homogeneous differential equations under a projective transformation. He finds first the induced transformation of the coefficients of a system of n linear homogeneous differential equations (in n dependent variables) of the m-th order and then proceeds to calculate the invariants for the particular case of $m = 2$, $n = 2$. This calculation, as presented there, requires the solution of a complete system, a process which is quite laborious even for the special case considered by Wilczynski. This method becomes more difficult if we try to calculate the invariants in the case $m = 2$, $n = 3$; and the calculations become more and more involved as the number of variables becomes larger.

It is our purpose to present an easier and more elegant solution of the problem of finding the invariants of a system of n linear homogeneous differential equations of the second order under a linear transformation. Consider the system

$$y_i'' + \sum_{j=1}^{n} L_{ij}(x)y_j' + \sum_{j=1}^{n} M_{ij}(x)y_j = 0 \qquad (i = 1, 2, \cdots, n),$$

where $y_i' = dy_i/dx$, $y_i'' = d^2y_i/dx^2$, under the transformation

$$y_i = \eta_i + \sum_{j=1}^{n} K_{ij}(x)\eta_j \qquad (i = 1, 2, \cdots, n).$$

We find in this paper the invariants which are functions of the arguments L_{ij}, $L_{ij}' = dL_{ij}/dx$, M_{ij} and also the invariants which are functions of the arguments L_{ij}, L_{ij}', L_{ij}'', M_{ij} and M_{ij}'.

Just as Wilczynski applied the theory for the case $n = 2$ to the study of ruled surfaces, we could also apply the results given in this paper to the study of a certain type of projective configuration represented by the system of differential equations here considered. However, this will be reserved for another work.

1. **Notation.** For the sake of simplicity we shall find it convenient to introduce the following permanent notations.

The small letters, for example y, η, shall represent vectors, that is, they will stand respectively for y_i, η_i $(i = 1, 2, \cdots, n)$. The letters x and a shall

Received August 2, 1939.

always stand for a single variable over the respective ranges $x_1 \leqq x \leqq x_2$, $-\infty < a < \infty$. The capital letters $L, M, G, U, \mathfrak{K}$, etc., will denote the matrices with elements $L_{ij}, M_{ij}, G_{ij}, U_{ij}, \mathfrak{K}_{ij}$ $(i, j = 1, 2, \cdots, n)$. The transpose of the matrix L will be denoted by $\bar{L}$.

It will be found convenient to introduce the notation $LM - ML \equiv [L, M]$. The trace of a matrix L will be denoted by tr L.

In general subscripts will be omitted in this paper. Only in the places where their presence will make for increased clarity in the expressions will subscripts be used. In these places the subscripts i and j, unless otherwise specified, will have the range of values $1, 2, \cdots, n$. All primed letters will denote differentiation with respect to the variable x.

2. **The equation and the transformation.** Consider the system of linear homogeneous differential equations of the second order

$$y'' + Ly' + My = 0, \tag{2.1}$$

where we assume the elements of the matrices L and M to be continuous functions of x possessing derivatives of all orders on the interval $x_1 \leqq x \leqq x_2$. The most general point transformation affecting only the dependent variables, which converts the system (2.1) into another of the same form and order, is given[1] by

$$y = \eta + K\eta, \tag{2.2}$$

where the elements of the matrix K are arbitrary continuous functions having derivatives of all orders on $x_1 \leqq x \leqq x_2$, and such that the determinant $|I + K|$ is not identically zero.

In order to obtain the infinitesimal transformations induced on L and M we now regard K as depending not only on the variable x but also on the single parameter a. We thus obtain from (2.1) and (2.2) the infinitesimal transformations

$$\frac{\partial L}{\partial a} = 2\mathfrak{K}' + [L, \mathfrak{K}], \qquad \frac{\partial M}{\partial a} = \mathfrak{K}'' + L\mathfrak{K}' + [M, \mathfrak{K}], \tag{2.3}$$

where the elements of the matrix $\mathfrak{K}$ are arbitrary continuous functions of x having derivatives of all orders. We shall also need the infinitesimal transformations of L', L'', M' which are found to be given by

$$\frac{\partial L'}{\partial a} = 2\mathfrak{K}'' + [L', \mathfrak{K}] + [L, \mathfrak{K}'], \tag{2.4}$$

$$\frac{\partial L''}{\partial a} = 2\mathfrak{K}''' + [L'', \mathfrak{K}] + 2[L', \mathfrak{K}'] + [L, \mathfrak{K}''], \tag{2.5}$$

[1] Wilczynski, loc. cit., Chapter I.

$$\frac{\partial M'}{\partial a} = \mathfrak{K}''' + L'\mathfrak{K}' + L\mathfrak{K}'' + [M', \mathfrak{K}] + [M, \mathfrak{K}']. \tag{2.6}$$

We may now exhibit the invariants which are functions of L, L', L'', M and M'. To do this define G and U by

$$G = 2L' - 4M + L^2, \qquad U = 2G' + [L, G], \tag{2.7}$$

and consider the matrices $G^l U^k$ $(l, k = 0, 1, \cdots)$. We will prove now that the trace tr $(G^l U^k)$ is an invariant. In order to do that we find the following infinitesimal transformations

$$\frac{\partial G^l}{\partial a} = [G^l, \mathfrak{K}], \qquad \frac{\partial U^k}{\partial a} = [U^k, \mathfrak{K}], \tag{2.8}$$

which are derived by the use of the equations (2.3), (2.4), (2.5) and (2.6). From (2.8) it follows at once that ∂ tr $(G^l U^k)/\partial a = 0$, and this proves that tr $(G^l U^k)$ is an invariant under the transformations induced on the coefficients L and M of (2.1) by the projective transformation (2.2).

3. **The invariants depending upon** L, L' **and** M. If the scalar function f is an invariant depending only upon the arguments L_{ij}, L'_{ij} and M_{ij}, the expression

$$\frac{\partial f}{\partial a} = \mathrm{tr}\left(\bar{\mathfrak{L}}\frac{\partial L}{\partial a} + \bar{\mathfrak{L}}'\frac{\partial L'}{\partial a} + \bar{\mathfrak{M}}\frac{\partial M}{\partial a}\right),$$

where

$$\frac{\partial f}{\partial L_{ij}} = \mathfrak{L}_{ij}, \qquad \frac{\partial f}{\partial L'_{ij}} = \mathfrak{L}'_{ij}, \qquad \frac{\partial f}{\partial M_{ij}} = \mathfrak{M}_{ij}$$

must vanish for all values of the arbitrary functions $\mathfrak{K}_{ij}$, $\mathfrak{K}'_{ij}$, $\mathfrak{K}''_{ij}$. Substituting the values from (2.3) and (2.4) and equating to zero the coefficients of these arbitrary functions, we obtain a system of $3n^2$ partial differential equations for f. By the general theory of Lie we know that this set of equations will form a complete system, and that any solution of it is an invariant, and conversely every invariant depending upon L_{ij}, L'_{ij}, M_{ij} is a solution of this complete system. This system is given by

$$\begin{aligned} &\text{(a)} \quad 2\mathfrak{L}' + \mathfrak{M} = 0, \\ &\text{(b)} \quad 2\mathfrak{L} + [\bar{L}, \mathfrak{L}'] + \bar{L}\mathfrak{M} = 0, \\ &\text{(c)} \quad [\bar{L}, \mathfrak{L}] + [\bar{L}', \mathfrak{L}'] + [\bar{M}, \mathfrak{M}] = 0. \end{aligned} \tag{3.1}$$

This complete system contains $3n^2$ equations with $3n^2$ independent variables. We shall now show that at least $3n^2 - n$ equations of the equations (3.1) are independent.

The n^2 equations in (3.1a) tell us that L'_{ij} and M_{ij} can occur only in the combinations $L'_{ij} - 2M_{ij}$. Hence the n^2 equations of (3.1b) show that L_{ij},

L'_{ij} and M_{ij} can occur only in the combinations defined by G_{ij} in (2.7). Introducing G_{ij} as new variables into the equations (3.1c), we obtain the following complete system of n^2 partial differential equations

$$\bar{G}\frac{\partial f}{\partial G} - \frac{\partial f}{\partial G}\bar{G} = 0. \tag{3.2}$$

Any solution of (3.2) is an invariant, and conversely every invariant is a solution of (3.2).

Consider the n^2 homogeneous linear equations (3.2) in the n^2 unknowns $\partial f/\partial G_{ij}$. Order these equations in the sequence 11, 12, $\cdots$, $1n$, 21, 22, $\cdots$, $2n$, 31, $\cdots$, nn and order the unknowns in these equations in the same way. Let the matrix of coefficients in the resulting set of equations be R. Now if $G_{ij} = 0$ for $i \neq j$, the matrix R is such that every element out of the diagonal is zero, n of the elements in the diagonal are zero and the remaining $n^2 - n$ are $G_{ii} - G_{jj}$ $(i \neq j)$. Hence the matrix R has the rank at least equal to $n^2 - n$, so that there are at most n independent invariants.

From the preceding section it follows that the n expressions tr (G^l) $(l = 1, 2, \cdots, n)$ are invariants. We shall now show that these n invariants are functionally independent in the n^2 variables G_{ij}. To do this it is sufficient to show that they are independent when considered as functions of the n variables $G_{11}, G_{22}, \cdots, G_{nn}$ only. By direct differentiation we have

$$\frac{\partial \operatorname{tr}(G^l)}{\partial G_{ij}} = lG_{ji}^{(l-1)}.$$

Making use of this formula, we see that after setting $G_{ij} = 0$, when $i \neq j$, the Jacobian of the n expressions tr (G^l) $(l = 1, 2, \cdots, n)$ with respect to the n variables $G_{11}, G_{22}, \cdots, G_{nn}$ is equal to n times the Vandermonde determinant in these n variables. This proves that the expressions tr (G^l) $(l = 1, 2, \cdots, n)$ are n functionally independent invariants. Combining the results of this section, we have

THEOREM 3.1. *Let $y'' + Ly' + My = 0$ be a system of n linear homogeneous differential equations of the second order. Suppose that the variable y is transformed by means of the transformation*

$$y = \eta + K\eta. \tag{3.3}$$

Then if we define $G = 2L' - 4M + L^2$, the first n successive traces of G are a system of n independent functions which remain unchanged by the transformation (3.3), *and any invariant depending upon L_{ij}, L'_{ij}, M_{ij} may be expressed functionally in terms of these traces.*

4. **The number of independent invariants depending upon** L, L', L'', M **and** M'. In the remaining sections of this paper we shall find the invariants depending only upon the arguments L, L', L'', M and M'. Although the number of such invariants and the invariants themselves have been found in

§§2 and 4 for the general n, the independence of this system of invariants has so far been shown only for $n = 2, 3, 4$.

Proceeding exactly in the same way as at the beginning of the preceding section, we obtain the following complete system of partial differential equations:

$$(4.1)\quad \begin{aligned} &(a)\quad 2\mathfrak{L}'' + \mathfrak{M}' = 0,\\ &(b)\quad 2\mathfrak{L}' + \mathfrak{M} + [\bar{L}, \mathfrak{L}''] + \bar{L}\mathfrak{M}' = 0,\\ &(c)\quad 2\mathfrak{L} + [\bar{L}, \mathfrak{L}'] - 2\mathfrak{L}''\bar{L}' + \bar{L}\mathfrak{M} + [\bar{M}, \mathfrak{M}'] = 0,\\ &(d)\quad [\bar{L}, \mathfrak{L}] + [\bar{L}', \mathfrak{L}'] + [\bar{L}'', \mathfrak{L}''] + [\bar{M}, \mathfrak{M}] + [\bar{M}', \mathfrak{M}'] = 0. \end{aligned}$$

This system contains $4n^2$ equations in the $5n^2$ variables L, L', L'', M and M'. Every solution of (4.1) is an invariant and every invariant depending upon L, L', L'', M and M' is a solution of this complete system. Denoting the matrix of the left side of (4.1d) by X, we see that

$$(4.2)\qquad \operatorname{tr} X = 0.$$

In the sequel it will be shown that $4n^2 - 1$ equations of (4.1) are independent.

From equations (4.1a) and (4.1b) it follows that the invariants are functions of the $3n^2$ arguments L, G and G'. After some calculation it may be shown from (4.1c) that our invariants are functions of G and U as defined in (2.7). Now introducing G and U as new independent variables into (4.1d), we obtain the following complete system of n^2 partial differential equations

$$(4.3)\qquad \bar{U}\frac{\partial f}{\partial U} - \frac{\partial f}{\partial U}\bar{U} + \bar{G}\frac{\partial f}{\partial G} - \frac{\partial f}{\partial G}\bar{G} = 0$$

in the $2n^2$ independent variables G and U. Any solution $f(G_{11}, G_{12}, \cdots, G_{nn}; U_{11}, U_{12}, \cdots, U_{nn})$ of (4.3) is an invariant, and conversely every invariant depending upon L, L', L'', M and M' is a solution of (4.3).

We shall now show that $n^2 - 1$ equations of (4.3) are independent, leaving (4.2) as the only relation among the equations (4.3). To do this it is sufficient to prove that the matrix of the coefficients of (4.3) is of rank $n^2 - 1$. This matrix may be described in the following manner. It consists of n^2 rows and $2n^2$ columns, n^2 columns of which have as their elements the same expressions in U_{ij} as the remaining n^2 columns have in G_{ij}. The variables $U_{11}, U_{22}, \cdots, U_{nn}$ appear only in the binomial form $U_{ii} - U_{jj}$ in the column of the coefficients of $\partial f/\partial U_{ij}$. Out of this matrix pick a determinant of order $n^2 - 1$ in the following way: Omit the row of coefficients corresponding to $i = n$, $j = n$; out of the n^2 columns associated with the coefficients of $\partial f/\partial U_{ij}$ take the $n^2 - n$ columns when $i \neq j$, and out of the n^2 columns associated with the coefficients of $\partial f/\partial G_{ij}$ take the $n - 1$ columns of coefficients of $\partial f/\partial G_{1p}$ $(p = 2, 3, \cdots, n)$. We find that this determinant, except possibly for sign, is equal to

$$G_{12}G_{13}\cdots G_{1n}\prod_{\substack{i,j=1\\ i<j}}^{n}(U_{ii} - U_{jj})^2 \neq 0.$$

Since the complete system (4.3) consists of $n^2 - 1$ independent equations, it follows that there are exactly $n^2 + 1$ functionally independent invariants involving the arguments L, L', L'', M and M'.

5. **Several lemmas.** It was shown in §2 that tr $(G^l U^k)$ are invariants involving L, L', L'', M and M'. In the preceding section it was shown that the number of all such functionally independent invariants is $n^2 + 1$. Our problem of finding the complete system of invariants involving the arguments L, L', L'', M and M' will then be solved if we can exhibit $n^2 + 1$ functionally independent traces tr $(G^l U^k)$. We consider then the following $n^2 + 1$ traces

$$\text{(5.1)} \quad \operatorname{tr}(G^l U^k) \quad (l = 1, 2, \cdots, n; k = 0, 1, \cdots, n-1); \quad \operatorname{tr}(U^n),$$

which we desire to prove to be functionally independent.

For these considerations it will be found convenient to introduce in this section some preliminary matters concerning the derivatives of the traces of composite matrices as well as properties of certain determinants which arise here.

We first have[2]

$$\text{(5.2)} \qquad \frac{\partial \operatorname{tr}(G^l U^k)}{\partial U_{ij}} = \sum_{m=1}^{n} \sum_{s=1}^{n} \sum_{p=1}^{k} U_{jm}^{(k-p)} G_{ms}^{(l)} U_{si}^{(p-1)}.$$

Consider an arbitrary n by n square matrix M_{ij}, the determinant of which is denoted by m. As usual $M_{ij}^{(k)}$ stands for the k-th power of the given matrix M_{ij}. Consider now determinants of the form

$$\text{(5.3)} \qquad | M_{r_1 s_1}^{(i)} M_{r_2 s_2}^{(i)} \cdots M_{r_n s_n}^{(i)} |,$$

where $r_1, r_2, \cdots, r_n, s_1, s_2, \cdots, s_n$ are arbitrary integers on the range 1 to n, equal or unequal, and where we are denoting a determinant by writing its i-th row. We call a determinant of the type (5.3) *a generalized determinant of Vandermonde.*

The generalized determinants of Vandermonde have many interesting properties, two of which will be stated here. We have

LEMMA 5.1. $| M_{r_1 s_1}^{(i)} M_{r_2 s_2}^{(i)} \cdots M_{r_n s_n}^{(i)} | = m \, | M_{r_1 s_1}^{(i-1)} M_{r_2 s_2}^{(i-1)} \cdots M_{r_n s_n}^{(i-1)} |.$

The proof of this result rests upon the fact that the matrix M satisfies its characteristic equation. From Lemma 5.1 follows

LEMMA 5.2. *If* $r_1 \neq s_1, r_2 \neq s_2, \cdots, r_n \neq s_n$ *simultaneously, then the value of the determinant* (5.3) *is zero.*

6. **The independence of the invariants.** In the present section we prove the independence of the invariants (5.1), for the cases $n = 2, 3, 4$.

We first prove the independence of the invariants in the case $n = 3$. Set up the Jacobian of the ten invariants (5.1) with respect to the variables G_{11}, $G_{12}, \cdots, G_{33}, U_{11}, U_{12}, \cdots, U_{33}$ and make use of (5.2). In order to show that this ten by eighteen matrix is of rank ten, pick three out of the first nine

[2] H. W. Turnbull, *A matrix form of Taylor's theorem*, Proceedings of the Edinburgh Mathematical Society, (2), vol. 2(1930–1931), p. 39, Theorem IV.

columns and seven out of the last nine columns. Setting $U_{ij} = 0$ when $i \neq j$ and $U_{11} = -U_{22} = 1$ and $U_{33} = 0$, we are thus led to

$$(6.1) \quad | G_{33}^{(i)} \quad G_{21}^{(i)} \quad G_{12}^{(i)} | \cdot \begin{vmatrix} G_{11}^{(i)} & G_{21}^{(i)} & G_{31}^{(i)} & G_{12}^{(i)} & G_{22}^{(i)} & G_{32}^{(i)} & G_{33}^{(i)} \\ 2G_{11}^{(i)} & 0 & G_{31}^{(i)} & 0 & -2G_{22}^{(i)} & -G_{32}^{(i)} & 0 \\ 3 & 0 & 0 & 0 & 3 & 0 & 0 \end{vmatrix},$$

where we are employing the notation introduced in the preceding section for the representation of generalized determinants of Vandermonde. By the Laplace development (6.1) is equal to $-6 \mid G_{33}^{(i)} \; G_{21}^{(i)} \; G_{12}^{(i)} \mid^2 [\mid G_{22}^{(i)} \; G_{31}^{(i)} \; G_{32}^{(i)} \mid + \mid G_{11}^{(i)} \; G_{31}^{(i)} \; G_{32}^{(i)} \mid]$. Denoting by g the determinant of the matrix G and applying Lemma 5.1, we see that (6.1) reduces finally to $-12g^2 \mid G_{33}^{(i-1)} \; G_{21}^{(i-1)} \; G_{12}^{(i-1)} \mid^2 \cdot \mid G_{22}^{(i-1)} \; G_{31}^{(i-1)} \; G_{32}^{(i-1)} \mid$, which can be readily verified not to be identically zero.

The proof of the functional independence of the five invariants (5.1) for the case $n = 2$, goes through in an exactly analogous fashion by taking the set of values $U_{ij} = 0$, $i \neq j$, $U_{11} = 1$ and $U_{22} = 0$.

Similarly we may prove the independence of the 17 invariants (5.1) in the case $n = 4$. The calculations are lengthy and will be omitted. We proceed as in the case $n = 3$: let $U_{ij} = 0$, $i \neq j$, $U_{11} = -U_{22} = 1$, $U_{33} = 2$, $U_{44} = 0$ and make use of Lemmas 5.1 and 5.2.

We have thus our final result:

THEOREM 6.1. *Let $y'' + Ly' + My = 0$ be a system of n linear homogeneous differential equations of the second order, for $n = 2, 3, 4$. Suppose that the variable y is transformed by means of*

$$(6.2) \qquad y = \eta + K\eta.$$

Then if we define $G = 2L' - 4M + L^2$, $U = 2G' + [L, G]$, the traces tr $(G^l U^k)$ $(l = 1, 2, \cdots, n; k = 0, 1, \cdots, n - 1)$, tr (U^n), *for $n = 2, 3, 4$, are a system of $n^2 + 1$ independent functions of G_{ij} and U_{ij} which remain unchanged by the transformation* (6.2), *and any invariant depending upon L_{ij}, L'_{ij}, L''_{ij}, M_{ij} and M'_{ij} may be expressed functionally in terms of these $n^2 + 1$ traces.*

As mentioned in the introduction to this paper, this theory has been worked out by Wilczynski (loc. cit., pp. 95–100), for the case $n = 2$. It will be of interest to compare the invariants obtained by Wilczynski with those obtained here for the case $n = 2$. The corresponding invariants of Wilczynski are in his notation

$$I = G_{11} + G_{22}, \qquad J = G_{11}G_{22} - G_{12}G_{21}, \qquad I' = G'_{11} + G'_{22},$$
$$J' = G'_{11}G_{22} + G_{11}G'_{22} - G'_{12}G_{21} - G_{12}G'_{21}, \qquad K = U_{11}U_{22} - U_{12}U_{21}.$$

We may easily verify that

$$\mathrm{tr}\,(G) = I, \qquad \mathrm{tr}\,(G^2) = I^2 - 2J, \qquad \mathrm{tr}\,(GU) = 2(II' - J'),$$
$$\mathrm{tr}\,(G^2U) = 2(I^2I' - IJ' - I'J), \qquad \mathrm{tr}\,(U^2) = 4I'^2 - 2K.$$

UNIVERSITY OF CINCINNATI.

ENTIRE FUNCTIONS BOUNDED ON A LINE

By R. P. Boas, Jr.

In the first part of this paper I give a new proof of Miss Cartwright's theorem that an entire function $f(z)$, of exponential type $k < \pi$, bounded at the integers, is bounded on the real axis.[1] It resembles Macintyre's proof in depending on an absolutely convergent interpolation series for $f(z)$; my interpolation series is simpler in appearance than Macintyre's, and is obtained in a different way. I use the principle that a function of exponential type with any pretensions to boundedness on the real axis is "nearly" a finite trigonometric integral; for such an integral, the interpolation series is easily obtained by a method[2] which seems to have been introduced by J. M. Whittaker.[3]

I then use the Paley-Wiener theory of "non-harmonic Fourier series"[4] to obtain an interpolation series involving $f(\lambda_n)$ instead of $f(n)$, where the λ_n are real numbers satisfying $|\lambda_n - n| \leqq L < 1/(2\pi^2)$ $(n = 0, \pm 1, \pm 2, \cdots)$. This allows me to replace n by λ_n in Miss Cartwright's theorem.

The generalized theorem is less interesting for itself than for its applications. Suppose that $f(z)$ is of exponential type $k < \pi$, that $\varphi(t)$ is a non-negative, non-decreasing and unbounded function in $t \geqq 0$, and that for some $\delta > 0$ (with real x)

$$\sup_{-\infty < n < \infty} \int_n^{n+\delta} \varphi(|f(x)|)\, dx < \infty. \tag{1}$$

Then it is obvious that there are numbers λ_n and A such that $|\lambda_n - n| < \frac{1}{20} < 1/(2\pi^2)$ and $|f(\lambda_n)| \leqq A$ for $n = 0, \pm 1, \pm 2, \cdots$. Consequently, (1) implies that $f(x)$ is bounded for real x. In particular, then, a function of (any) exponential type satisfying such a condition as

$$\int_{-\infty}^{\infty} |f(x)|^p\, dx < \infty, \text{ with } p > 0,$$

or

$$\sup_{-\infty < x < \infty} \int_x^{x+1} \log^+ |f(t)|\, dt < \infty,$$

Received August 17, 1939. The author is a National Research Fellow.

[1] Cartwright [1], Pfluger [14], Macintyre [12]. Numbers in brackets refer to the list of references at the end of the paper.

[2] Cf. Pólya and Szegö [16], vol. 1, p. VI, line 35.

[3] See [20], p. 67. Cf. also Gelfond [6].

[4] Paley and Wiener [13], Chapter 7. I have not been able to use the more refined results of Levinson [10, 11].

is bounded on the real axis.[5] A further consequence is that if $f(z)$ is of order one and minimum type, and satisfies (1), it is a constant. This includes results of several authors[6] in which the hypothesis is of the form

$$\int_{-\infty}^{\infty} \varphi(|f(x)|)\,dx < \infty, \tag{2}$$

with various special functions φ. (The same result follows also from Levinson's theorem [9] that the boundedness of $f(\lambda_n)$ implies that $f(x) = O(|x|^A)$, $|x| \to \infty$, for some A, provided that $|\lambda_n - n| < a$, and $|\lambda_n - \lambda_m| \geqq \delta > 0$ for $n \neq m$.) The gain in generality from replacing (2) by (1) is considerable, since (1) fails only if $f(x)$ has too many large values, while (2) may fail merely because $f(x)$ has too many values bounded from zero.

I also apply my interpolation formulas to obtain and generalize some of the results of Plancherel and Pólya [15] of the types

$$\int_{-\infty}^{\infty} |f(x)|^p\,dx \leqq A(k, p) \sum_{n=-\infty}^{\infty} |f(n)|^p,$$

$$\sum_{n=-\infty}^{\infty} |f(x_n)|^p \leqq A(k, p) \int_{-\infty}^{\infty} |f(x)|^p\,dx$$

(see Theorems 3 and 7).

As an extra application of a uniqueness theorem which I have to prove, I give (in §9) a new proof of a theorem of Levinson [11] on the closure of $\{e^{i\lambda_n x}\}$, and a discussion of the analogue for such sets of functions of the Weierstrass approximation theorem.

I use throughout the convention that A, $A(p)$, etc., denote numbers depending only on the indicated parameters, but in general varying from one occurrence to another. Numbers denoted by other letters preserve their identities until explicitly redefined.

1. Theorem 1. *If*

$$f(z) = \int_{-k}^{k} e^{izt}\,d\alpha(t) \qquad (0 < k < \pi), \tag{1.1}$$

[5] For the first condition, see Plancherel and Pólya [15], p. 124. **Added in proof:** A. C. Schaeffer has communicated to me a very simple proof, obtained independently, of the boundedness of a function of exponential type satisfying (2) with any non-negative $\varphi(x)$ whose inferior limit as $x \to \infty$ is positive. His method of proof does not appear to be available under more general hypotheses such as (1).

[6] Paley and Wiener [13], p. 13; Siegel [17], p. 323; Iyer [8], p. 296; Plancherel and Pólya [15], p. 125; Boas [2], p. 278. To prove S. Bernstein's theorem that an entire function of order one and minimum type, bounded on the real axis, is a constant, one may (for example) apply Bernstein's inequality for the derivative of a function of exponential type (see Pólya and Szegö [16], problem IV 201, vol. 2, p. 35).

where $\alpha(t)$ *is a function of bounded variation on* $(-k, k)$, *then*

$$f(z) = \frac{\sin \pi z}{\pi(\pi - k)} \sum_{n=-\infty}^{\infty} \frac{(-1)^n f(n) \sin \{(\pi - k)(z - n)\}}{(z - n)^2}. \tag{1.2}$$

We define $Q(t, z)$ by

$$Q(t, z) = \begin{cases} e^{izt} & (-k \leqq t \leqq k), \\ e^{izt} + e^{izt}(e^{-2\pi iz} - 1) \dfrac{t - k}{2(\pi - k)} & (k \leqq t \leqq 2\pi - k). \end{cases}$$

If $\beta(t) = \alpha(t)$ on $(-k, k)$, and $\beta(t) = \alpha(k)$ on $(k, 2\pi - k)$, we then have

$$f(z) = \int_{-k}^{2\pi-k} Q(t, z)\, d\beta(t). \tag{1.3}$$

We now form the Fourier series of $Q(t, z)$, as a function of t. We find that the coefficient of e^{int} in this Fourier series is

$$c_n(z) = \frac{(-1)^n \sin \pi z \sin \{(\pi - k)(z - n)\}}{\pi(\pi - k)(z - n)^2}.$$

Since the Fourier series is uniformly convergent, we may substitute it for its sum in (1.3) and integrate term by term. Noting that

$$\int_{-k}^{2\pi-k} e^{int}\, d\beta(t) = f(n) \qquad (n = 0, \pm 1, \pm 2, \cdots),$$

we obtain (1.2).

We now extend Theorem 1 to more general functions.

THEOREM 2. *If* $f(z)$ *is an entire function of exponential type* $k < \pi$, *and if*

$$|f(n)| < K \qquad (n = 0, \pm 1, \pm 2, \cdots), \tag{1.4}$$

then (1.2) *holds.*

We define numbers b_n by

$$b_0 = f'(0), \qquad b_n = \frac{f(n) - f(0)}{n}.$$

Then by the Riesz-Fischer theorem there is a function $h(t) \in L^2(-\pi, \pi)$ such that

$$b_n = \int_{-\pi}^{\pi} e^{int} h(t)\, dt \qquad (n = 0, \pm 1, \pm 2, \cdots).$$

Let $H(z) = \int_{-\pi}^{\pi} e^{izt} h(t)\, dt$. Simple calculation shows that

$$|H(re^{i\theta})| \leqq \epsilon(r) e^{\pi r}, \tag{1.5}$$

where $\epsilon(r) = o(r)$ as $r \to \infty$. The function $H_1(z) = H(z) - z^{-1}\{f(z) - f(0)\}$ also satisfies (1.5); and $H_1(n) = 0$ for $n = 0, \pm 1, \pm 2, \cdots$. By a theorem of

Valiron,[7] $H_1(z) = c \sin \pi z$. It is clear that $c = 0$ since, as $|y| \to \infty$, $H_1(iy) = o(e^{\pi|y|})$ and $\sin \pi iy = \Omega(e^{\pi|y|})$. Thus $H(z) = z^{-1}\{f(z) - f(0)\}$. $H(z)$ is therefore of type k, and consequently $h(t)$ is almost everywhere zero on $(-\pi, -k)$ and on (k, π). That is, we have

$$f(z) = f(0) + z \int_{-k}^{k} e^{izt} h(t)\, dt, \qquad h(t) \in L^2(-k, k).$$

There are functions $h_\nu(t)$, each of bounded variation on $(-k, k)$, such that $h_\nu(k) = h_\nu(-k) = 0$, and[8] $h(t) = \underset{\nu\to\infty}{\text{l.i.m.}}\, h_\nu(t)$. If

$$f_\nu(z) = f(0) + z \int_{-k}^{k} e^{izt} h_\nu(t)\, dt,$$

we have

$$f(z) - f_\nu(z) = z \int_{-k}^{k} e^{izt}\{h(t) - h_\nu(t)\}\, dt \to 0 \tag{1.6}$$

as $\nu \to \infty$, for each z. We note that

$$\begin{aligned} &|f_\nu'(0)|^2 + \sum_{n\neq 0} \left| \frac{f_\nu(n) - f(0)}{n} \right|^2 \\ &\qquad = 2\pi \int_{-k}^{k} |h_\nu(t)|^2\, dt \to 2\pi \int_{-k}^{k} |h(t)|^2\, dt \quad (\nu \to \infty). \end{aligned} \tag{1.7}$$

Now

$$f_\nu(z) = f(0) + i \int_{-k}^{k} e^{izt}\, dh_\nu(t) = \int_{-k}^{k} e^{izt}\, dh_\nu^*(t),$$

where $h_\nu^*(t) = ih_\nu(t) + f(0) \operatorname{sgn} t$. By Theorem 1,

$$f_\nu(z) = \frac{\sin \pi z}{\pi(\pi - k)} \sum_{n=-\infty}^{\infty} \frac{(-1)^n f_\nu(n) \sin\{(\pi - k)(z - n)\}}{(z - n)^2}.$$

Write

$$g(z) = \frac{\sin \pi z}{\pi(\pi - k)} \sum_{n=-\infty}^{\infty} \frac{(-1)^n f(n) \sin\{(\pi - k)(z - n)\}}{(z - n)^2}.$$

Then if z is fixed and $N > 2|z|$,

$$\begin{aligned} \pi(\pi - k)|g(z) - f_\nu(z)| &\leqq \sum_{n=-N}^{N} |f(n) - f_\nu(n)| \frac{|\sin \pi z \sin\{(\pi - k)(z - n)\}|}{|z - n|^2} \\ &\qquad + 4e^{(2\pi - k)|y|}\left\{ \sum_{|n|\geqq N} \frac{|f_\nu(n)|}{n^2} + \sum_{|n|\geqq N} \frac{|f(n)|}{n^2} \right\} \\ &= S_1 + 4e^{(2\pi - k)|y|}(S_2 + S_3), \end{aligned}$$

say.

[7] Valiron [19], p. 204. Lemma 7, below, contains Valiron's theorem as a special case.
[8] "l.i.m." means "limit in the mean of order 2".

Evidently $S_3 \to 0$ as $N \to \infty$. The same is true of S_2, uniformly with respect to ν, since

$$\left\{\sum_{|n|\geqq N} \frac{|f_\nu(n)|}{n^2}\right\}^2 \leqq \sum_{|n|\geqq N} \frac{|f_\nu(n)|^2}{n^2} \sum_{|n|\geqq N} \frac{1}{n^2};$$

the first sum on the right is uniformly bounded, by (1.7), and the second approaches zero. For fixed N, $S_1 \to 0$ as $\nu \to \infty$, by (1.6).

Since $f_\nu(z) \to f(z)$ and $|g(z) - f_\nu(z)| \to 0$, $g(z) \equiv f(z)$; that is, $f(z)$ satisfies (1.2).

Miss Cartwright's theorem is an immediate consequence. For, if $|f(n)| \leqq K$, $f(z)$ is represented by (1.2), and we obtain

$$\begin{aligned} |f(x)| &\leqq \frac{K}{\pi(\pi - k)}\left\{\sum_{n=-\infty}^{[x]-1} + \sum_{n=[x]+2}^{\infty}\right\}\frac{1}{(x-n)^2} + 2K \\ &\leqq B(k)K, \end{aligned}$$

with[9]

$$B(k) = 2 + \frac{\pi}{3(\pi - k)}.$$

2. We can use Theorem 2 to establish a theorem which includes some results of Plancherel and Pólya connecting $\sum |f(n)|^p$ and $\int |f(x)|^p\,dx$.

THEOREM 3. *Let $\beta(x)$ be a non-decreasing function defined on $(-\infty, \infty)$ such that $\beta(x+1) - \beta(x) \leqq B < \infty$. Let $\gamma(t)$ be a non-decreasing function defined on $(-\delta, \delta)$, $\delta > 0$. If $f(z)$ is an entire function of exponential type $k < \pi$, and $p \geqq 1$, then*

$$\int_{-\infty}^{\infty} |f(x)|^p\,d\beta(x) \leqq \frac{B\{C(k,\delta)\}^p}{\gamma(\delta) - \gamma(-\delta)} \sum_{n=-\infty}^{\infty} \int_{-\delta}^{\delta} |f(n+t)|^p\,d\gamma(t). \tag{2.1}$$

We may take

$$C(k, \delta) = 2\left(\frac{1}{\pi(\pi - k)} + \delta + 3\right).$$

Before proving Theorem 3, we notice two of its interesting special cases.

(i) Let $\beta(x) = x$, $\gamma(t) = \operatorname{sgn} t$. The result is[10]

$$\int_{-\infty}^{\infty} |f(x)|^p\,dx \leqq \{C(k, 0)\}^p \sum_{-\infty}^{\infty} |f(n)|^p. \tag{2.2}$$

[9] Cf. Macintyre [12], p. 5, where a rather more elaborate $B(k)$ is obtained.

[10] Plancherel and Pólya [15], p. 120.

(ii) Let $\gamma(t) = t$, $\delta = \frac{1}{2}$; let $\beta(x)$ be a step function with unit jumps at points x_n such that no interval of unit length contains more than B of the x_n's. Then

$$\begin{aligned}\sum_{n=-\infty}^{\infty} |f(x_n)|^p &\leqq B\{C(k, \tfrac{1}{2})\}^p \sum_{n=-\infty}^{\infty} \int_{-\frac{1}{2}}^{\frac{1}{2}} |f(n+t)|^p\,dt \\ &\leqq B\{C(k, \tfrac{1}{2})\}^p \int_{-\infty}^{\infty} |f(x)|^p\,dx.\end{aligned} \tag{2.3}$$

The restriction to functions of type less than π can be removed. For, if $f(z)$ is of type c, $g(z) = f(\frac{1}{2}\pi z/c)$ is of type $\frac{1}{2}\pi$, and we can apply (2.3) to $g(z)$ with points $x_n^* = 2cx_n/\pi$. We obtain

$$\sum_{n=-\infty}^{\infty} |f(x_n)|^p \leqq K(B, c, p) \int_{-\infty}^{\infty} |f(x)|^p\,dx, \tag{2.4}$$

with

$$K(B, c, p) = \frac{2Bc}{\pi}\left(\left[\frac{\pi}{2c}\right] + 1\right)\left\{8\left(\frac{1}{\pi^2} + 2\right)\right\}^p.$$

A similar result has been established by Plancherel and Pólya,[11] under the condition $|x_n - x_m| \geqq d > 0$ $(n \neq m)$; they obtain

$$K(B, c, p) = \frac{8}{\pi d^2 pc}\,(e^{\frac{1}{2}pcd} - 1).$$

We now prove Theorem 3. We suppose that the right side of (2.1) is finite. This implies that $\sum_{n=-\infty}^{\infty} |f(n+t)|^p$ converges, except perhaps for a set of measure zero with respect to $\gamma(t)$. Consequently, for t not in this set, the numbers $f(n+t)$ are bounded with respect to n, and we can apply Theorem 2 to the function $f(z+n+t)$ for any integer n. Writing (2.1) with $z = x - t$, $|x| \leqq \frac{1}{2}$, and $|t| \leqq \delta$, we have

$$|f(x+n)|^p \leqq \left\{\sum_{\nu=-\infty}^{\infty} T_\nu\,|f(\nu+n+t)|\right\}^p,$$

where $T_\nu = 1$ for $[\frac{1}{2} - \delta] - 2 \leqq \nu \leqq [\frac{1}{2} + \delta] + 2$, and

$$T_\nu = \frac{1}{\pi(\pi - k)\{(\frac{1}{2} + \delta)\operatorname{sgn}\nu - \nu\}^2}$$

for other values of ν. Integrating this inequality over $(-\delta, \delta)$ with respect to $\gamma(t)$, and applying Minkowski's inequality,[12] we obtain

$$\{\gamma(\delta) - \gamma(-\delta)\}\,|f(x+n)|^p \leqq \left\{\sum_{\nu=-\infty}^{\infty} T_\nu\left(\int_{-\delta}^{\delta} |f(\nu+n+t)|^p\,d\gamma(t)\right)^{1/p}\right\}^p.$$

[11] Plancherel and Pólya [15], p. 126.
[12] Hardy, Littlewood, and Pólya [7], p. 148.

Write

$$\mathfrak{M}_p[f(\nu)] = \left\{\frac{1}{\gamma(\delta) - \gamma(-\delta)} \int_{-\delta}^{\delta} |f(\nu + t)|^p \, d\gamma(t)\right\}^{1/p}.$$

Then

$$|f(x + n)|^p \leqq \left\{\sum_{\nu=-\infty}^{\infty} T_\nu \mathfrak{M}_p[f(\nu + n)]\right\}^p. \tag{2.5}$$

Since $\int_{-\frac{1}{2}}^{\frac{1}{2}} d\beta(x + n) \leqq B$, we have from (2.5)

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} |f(x + n)|^p \, d\beta(x + n) \leqq B\left\{\sum_{\nu=-\infty}^{\infty} T_\nu \mathfrak{M}_p[f(\nu + n)]\right\}^p. \tag{2.6}$$

Now

$$\sum_{n=-\infty}^{\infty} \int_{-\frac{1}{2}}^{\frac{1}{2}} |f(x + n)|^p \, d\beta(x + n) = \sum_{n=-\infty}^{\infty} \int_{n-\frac{1}{2}}^{n+\frac{1}{2}} |f(x)|^p \, d\beta(x) = \int_{-\infty}^{\infty} |f(x)|^p \, d\beta(x).$$

If we sum both sides of (2.6) with respect to n and apply Minkowski's inequality, we therefore have

$$\begin{aligned} \int_{-\infty}^{\infty} |f(x)|^p \, d\beta(x) &\leqq B\left\{\sum_{\nu=-\infty}^{\infty} T_\nu \left(\sum_{n=-\infty}^{\infty} \mathfrak{M}_p^p[f(\nu + n)]\right)^{1/p}\right\}^p \\ &\leqq B\left\{\sum_{\nu=-\infty}^{\infty} T_\nu\right\}^p \sum_{n=-\infty}^{\infty} \mathfrak{M}_p^p[f(n)]. \end{aligned} \tag{2.7}$$

A simple computation shows that

$$\sum_{\nu=-\infty}^{\infty} T_\nu \leqq \frac{2^p_*}{\pi^p(\pi - k)^p} + (2\delta + 6)^p;$$

substituting this in (2.7), we have the result of Theorem 3. All the formal inversions of limit operations are justified by Fubini's theorem.

Slight modification of the proof gives us (2.1) for $p > \frac{1}{2}$, but with an $A(k, \delta, p)$ instead of $\{C(k, \delta)\}^p$. It would be possible, by altering (1.2) to make it converge more rapidly, to obtain a similar result for any positive p; in the special cases (i) and (ii) this result has been obtained (by a different method) by Plancherel and Pólya. Both Theorem 3 and its generalization to all $p > 0$ are contained, in a rather less exact form, in Theorem 7.

3. To discuss functions bounded at points λ_n, we shall need several lemmas on the entire function

$$\psi(z) = (z - \lambda_0) \prod_1^{\infty} \left(1 - \frac{z}{\lambda_n}\right)\left(1 - \frac{z}{\lambda_{-n}}\right). \tag{3.1}$$

We suppose that

$$|\lambda_n - n| \leqq \delta \qquad (n = 0, \pm 1, \pm 2, \cdots), \tag{3.2}$$

where $0 < \delta \leqq \frac{1}{2}$. If $\delta = \frac{1}{2}$, we suppose that no two λ_n's are equal; this convention is adhered to throughout this section. It is convenient to write $\mu_n = -\lambda_{-n}$ $(n > 0)$ and $\psi(z) = (z - \lambda_0)\,\varphi(z)$, so that

$$\varphi(z) = \prod_1^\infty \left(1 - \frac{z}{\lambda_n}\right)\left(1 + \frac{z}{\mu_n}\right).$$

It is possible to estimate the product in (3.1) directly; however, we shall estimate it only for special values of z, and extend our estimate to general z by using Phragmén-Lindelöf theorems. This procedure is less straightforward, but avoids some unpleasant elementary calculations.

LEMMA 1. *For* $|y| \geqq 1$, $|\psi(iy)| \leqq A\,|y|^{2\delta}e^{\pi|y|}$.

We have

$$\left|\left(1 - \frac{iy}{\lambda_n}\right)\left(1 + \frac{iy}{\mu_n}\right)\right|^2 = 1 + \frac{y^4}{\lambda_n^2\mu_n^2} + y^2\left(\frac{1}{\mu_n^2} + \frac{1}{\lambda_n^2}\right) \leqq \left(1 + \frac{y^2}{(n-\delta)^2}\right)^2.$$

Hence

$$|\varphi(iy)| \leqq \prod_1^\infty \left(1 + \frac{y^2}{(n-\delta)^2}\right).$$

If $|y| \geqq 1$,

$$\begin{aligned}
\log|\varphi(iy)| - \log\frac{\sin \pi iy}{\pi iy} &\leqq \sum_{n=1}^\infty \left\{\log\left(1 + \frac{y^2}{(n-\delta)^2}\right) - \log\left(1 + \frac{y^2}{n^2}\right)\right\}\\
&\leqq y^2 \sum_{n=1}^\infty \left(\frac{1}{(n-\delta)^2} - \frac{1}{n^2}\right)\frac{n^2}{n^2+y^2}\\
&= 2y^2\delta \sum_{n=1}^\infty \frac{1}{(n-\delta)(n^2+y^2)} + y^2\delta^2 \sum_{n=1}^\infty \frac{1}{(n-\delta)^2(n^2+y^2)}\\
&\leqq A + 2\delta \sum_2^{[y]} \frac{1}{n-1} + 2\delta \sum_{[y]+1}^\infty \frac{|y|}{(n+|y|)(n-1)}\\
&\leqq A + 2\delta \log|y|.^{13}
\end{aligned}$$

Thus

$$|\varphi(iy)| \leqq A\,|y|^{2\delta-1}e^{\pi|y|},$$
$$|\psi(iy)| \leqq A\,|y|^{2\delta}e^{\pi|y|}.$$

LEMMA 2. *For* $-\infty < x < \infty$ *and* $|y| \leqq 1$, $|\psi(x+iy)| \leqq A$.

We first consider points $z = x \pm i$, with $2 \leqq x \leqq 3$.

[13] $y/(n^2+y^2) \leqq 1/(n+y)$ if $0 \leqq y \leqq n$. We recall our convention about A's, stated at the end of the introduction.

Then

$$|\psi(z)| = |z - \lambda_0| \prod_1^\infty \left|1 - \frac{z}{\lambda_n}\right|\left|1 + \frac{z}{\mu_n}\right| \leqq A \prod_4^\infty. \tag{3.3}$$

When $n \geqq 4$, $\lambda_n \geqq x + \frac{1}{2}$, and

$$\left|1 - \frac{x \pm i}{\lambda_n}\right|^2 = \left(1 - \frac{x}{\lambda_n}\right)^2 + \frac{1}{\lambda_n^2} \leqq \left|1 - \frac{x \pm i}{n+1}\right|^2, \tag{3.4}$$

since $(1 - x/\lambda)^2 + \lambda^{-2}$ is an increasing function of λ if $x > 2$ and $\lambda \geqq x + \frac{1}{2}$. The inequality

$$\left|1 + \frac{z}{\mu_n}\right| \leqq \left|1 + \frac{z}{n-1}\right| \tag{3.5}$$

is evident. Using (3.4) and (3.5) in (3.3), we have

$$\begin{aligned} |\psi(z)| &\leqq A \prod_4^\infty \left|1 + \frac{z}{n-1}\right|\left|1 - \frac{z}{n+1}\right| \\ &\leqq A \prod_3^\infty \left|1 - \frac{z^2}{n^2}\right| = A\left|\frac{\sin \pi z}{\pi z}\right| \frac{1}{|1 - z^2|\cdot|1 - \frac{1}{4}z^2|}, \end{aligned}$$

$$|\psi(x \pm i)| \leqq A \qquad (2 \leqq x \leqq 3). \tag{3.6}$$

When m is an integer, the functions $\psi_m(z) = \psi(z + m)$ are products like (3.1) with zeros $\lambda_n^{(m)} = \lambda_{n+m} - m$ satisfying $|\lambda_n^{(m)} - n| \leqq \delta$. Hence the $\psi_m(z)$ satisfy (3.6). But (3.6) for $\psi_{m-2}(z)$ states that $|\psi(x + m - 2 \pm i)| \leqq A$ for $2 \leqq x \leqq 3$, or that $|\psi(x \pm i)| \leqq A$ for $m \leqq x \leqq m + 1$. Since this is true for every m, $|\psi(x \pm i)| \leqq A$ for all x; since $\psi(z)$ is of order one, $|\psi(x + iy)| \leqq A$ for $|y| \leqq 1$, by the Phragmén-Lindelöf theorem for a strip.[14]

LEMMA 3.[15] *For all x and y, $|\psi(x + iy)| < A(1 + |y|)^{2\delta}e^{\pi|y|}$.*

We apply Lemma 1 to the functions $\psi_m(z)$ defined above, and deduce that

$$|\psi(m + iy)| \leqq A|y|^{2\delta}e^{\pi|y|}, \qquad |y| \geqq 1. \tag{3.7}$$

Now consider the functions

$$\omega_m(z) = \frac{\psi(z)}{(z - m + 1)^{2\delta} \exp[-i\pi z \operatorname{sgn} y]}$$

in the strips $m \leqq x \leqq m + 1$, $|y| \geqq 1$, taking the branch of $(z - m + 1)^{2\delta}$ which is $+1$ at $z = m$. For $|y| = 1$, $|\omega_m(z)| \leqq A$, by Lemma 2. For $|y| \geqq 1$ and $x = m$ or $x = m + 1$, $|\omega_m(z)| \leqq A$ by (3.7). By the same Phragmén-Lindelöf theorem as before, $|\omega_m(z)| \leqq A$ for $|y| \geqq 1$, $m \leqq x \leqq m + 1$. Consequently $|\psi(z)| \leqq A|y|^{2\delta}e^{\pi|y|}$ for $|y| \geqq 1$; this, combined with Lemma 2, gives the result of Lemma 3.

[14] Titchmarsh [18], p. 180.

[15] Levinson ([10], p. 925) proves that $|\psi(z)| < A(1 + |z|)e^{\pi|y|}$, if $\delta \leqq \frac{1}{4}$. This is not quite enough for our purposes.

LEMMA 4. *For* $|x| \leqq 1$ *and* $|y| \geqq 3$, $|\psi(x+iy)| \geqq A|y|^{-2}e^{\pi|y|}$.

This will be used in the proof of Lemma 6, which will supersede it.

When $|x| \leqq 1$ and $|y| \geqq |x| + \delta$,

$$\left|1 - \frac{z}{\lambda_n}\right|^2 \left|1 + \frac{z}{\mu_n}\right|^2 \geqq \left(1 + \frac{y^2 - (|x| + \delta)^2}{\lambda_n \mu_n}\right)^2 \geqq \left(1 + \frac{y^2 - 4}{(n+1)^2}\right)^2.$$

Consequently, when $|y| \geqq 3$,

$$\begin{aligned} |\psi(x+iy)| &\geqq A|y| \prod_1^\infty \left(1 + \frac{(|y| - 2)^2}{(n+1)^2}\right) \\ &= A|y| \frac{\sin\{\pi i(|y| - 2)\}}{\pi i(|y| - 2)\{1 + (|y| - 2)\}^2} \\ &\geqq A|y|^{-2}e^{\pi|y|}. \end{aligned}$$

LEMMA 5. *For* $|y| \geqq 3$, $|\psi(iy)| \geqq A|y|^{-2\delta}e^{\pi|y|}$.

We have

$$\left|1 - \frac{iy}{\lambda_n}\right|\left|1 + \frac{iy}{\mu_n}\right| \geqq 1 + \frac{y^2}{(n+\delta)^2};$$

$$\begin{aligned} \log\left|\frac{\pi iy\varphi(iy)}{\sin \pi iy}\right| &\geqq \sum_{n=1}^\infty \left\{\log\left(1 + \frac{y^2}{(n+\delta)^2}\right) - \log\left(1 + \frac{y^2}{n^2}\right)\right\} \\ &\geqq -\delta y^2 \sum_{n=1}^\infty \frac{2n + \delta}{n^2\{(n+\delta)^2 + y^2\}} \\ &\geqq -A - 2\delta y^2 \sum_1^\infty \frac{1}{n(n^2 + y^2)} \\ &\geqq -A - 2\delta \log|y|, \qquad |y| \geqq 3. \end{aligned}$$

Hence

$$|\varphi(iy)| \geqq A|y|^{-2\delta}\left|\frac{\sin \pi iy}{\pi iy}\right|,$$

$$|\psi(iy)| \geqq A|y|^{-2\delta}e^{\pi|y|}, \qquad |y| \geqq 3.$$

LEMMA 6.[16] *For* $|y| \geqq 3$, $|\psi(z)| \geqq A|y|^{-2\delta}e^{\pi|y|}$.

Consider again the functions $\psi_m(z) = \psi(z + m)$ $(m = 0, \pm 1, \pm 2, \cdots)$. By Lemma 5, applied to $\psi_m(z)$,

$$|\psi(m + iy)| \geqq A|y|^{-2\delta}e^{\pi|y|} \qquad (|y| \geqq 3;\ m = 0, \pm 1, \pm 2, \cdots). \tag{3.8}$$

[16] Levinson (loc. cit.) proves that $|\psi(z)| > A|y|(1 + |z|)^{-\frac{1}{2}}e^{\pi|y|}$, if $\delta \leqq \frac{1}{4}$. We could get along with this weaker inequality at the expense of additional complications in the proof of Lemma 9.

By Lemma 4,

$$|\psi(x \pm 3i)| \geqq \tfrac{1}{9} A e^{3\pi} = A > 0 \qquad (-\infty < x < \infty); \tag{3.9}$$

and, by Lemma 4 applied to $\psi_m(z)$,

$$|\psi(m + x + iy)| \geqq A\,|y|^{-2} e^{\pi|y|}, \qquad |x| \leqq 1. \tag{3.10}$$

The functions

$$\omega_m(z) = \frac{\exp[-i\pi z \operatorname{sgn} y]}{\psi(z)(z - m + 1)^{2\delta}},$$

where we take the branch of $(z - m + 1)^{2\delta}$ which is $+1$ at $z = m$, are regular in the strips $m \leqq x \leqq m + 1$, $|y| \geqq 3$; are bounded by an A on the boundaries of the strips, because of (3.8) and (3.9); and are $O(|y|^{2-2\delta})$ as $|y| \to \infty$ in each strip. Hence, by the Phragmén-Lindelöf theorem for a strip, $|\omega_m(z)| \leqq A$ throughout each strip, and the conclusion of Lemma 6 follows.

4. We use Lemma 6 to generalize Valiron's uniqueness theorem (which we used in §2); our proof follows the same lines as Valiron's.

LEMMA 7. *If $H(z)$ is an entire function such that*

$$|H(z)| \leqq K e^{\pi|z|}(1 + |y|)^{-2\delta} \qquad (0 \leqq \delta \leqq \tfrac{1}{2}), \tag{4.1}$$

and if $H(\lambda_n) = 0$ for $n = 0, \pm 1, \pm 2, \cdots$, where $|\lambda_n - n| \leqq \delta$ and, when $\delta = \frac{1}{2}$, no two λ_n's are equal, then $H(z)$ is a constant multiple of $\psi(z)$.

Because of Lemma 5, we obviously have the following corollary.

COROLLARY. *If the K in* (4.1) *is replaced by an $\epsilon(y)$ which is $o(1)$ as $|y| \to \infty$, then $H(z) \equiv 0$.*

To prove Lemma 7, we consider the entire function $\Phi(z) = H(z)/\psi(z)$. We have

$$\begin{aligned} |\Phi(re^{i\theta})| &\leqq B\,|\csc\theta|\exp\{\pi r(1 - |\sin\theta|)\} \\ &\leqq 2B\exp\{\pi r(1 - |\sin\theta|)\}, \qquad \tfrac{1}{2}\pi \geqq |\theta| \geqq \tfrac{1}{3}\pi, \end{aligned} \tag{4.2}$$

where B depends on K (and on ψ and H) but not on r and θ. This is a consequence of Lemma 6 and (4.1) for large r, and trivial for small r. We write, for $\frac{1}{3}\pi < \beta < \frac{1}{2}\pi$,

$$\Psi_\beta(z) = \Phi(z)\exp\{-\pi z(1 - \sin\beta)\sec\beta\}.$$

When $\arg z = \pm\beta$, (4.2) shows that

$$|\Psi_\beta(z)| \leqq 2\beta. \tag{4.3}$$

Hence, by the Phragmén-Lindelöf theorem for an angle,[17] (4.3) holds for $|\arg z| \leqq \beta$, so that

$$|\Phi(re^{i\theta})| \leqq 2B\exp\{\pi r(1 - \sin\beta)\sec\beta\cos\theta\} \tag{4.4}$$

[17] Titchmarsh [18], p. 177.

for $|\theta| \leqq \beta$. Since $(1 - \sin\beta)\sec\beta \downarrow 0$ as $\beta \uparrow \frac{1}{2}\pi$, (4.2) implies (4.4) also for $\frac{1}{2}\pi \geqq \theta \geqq \beta$. Since B is independent of β, we may let $\beta \to \frac{1}{2}\pi$ in (4.4); it follows that

$$|\Phi(re^{i\theta})| \leqq 2B, \qquad 0 \leqq |\theta| < \tfrac{1}{2}\pi,$$

and hence, since similar reasoning applies to $\Phi(-z)$, that $\Phi(z)$ is a constant.

5. We now collect the results which we shall need from the theory of non-harmonic Fourier series.

Lemma 8. *Suppose that the numbers λ_n are real, and that $|\lambda_n - n| \leqq L < 1/\pi^2$ $(n = 0, \pm1, \pm2, \cdots)$.*

(i)[18] *If $f(t) \in L^2(-\pi, \pi)$, there are numbers b_n such that*

$$f(t) = \underset{N\to\infty}{\text{l.i.m.}} \sum_{n=-N}^{N} b_n e^{i\lambda_n t}, \tag{5.1}$$

$$\sum_{n=-\infty}^{\infty} |b_n|^2 \leqq B(L) \int_{-\pi}^{\pi} |f(t)|^2 dt, \tag{5.2}$$

with $B(L) = (2\pi)^{-1}(1 - \pi L^{\frac{1}{2}})^{-2}$.

(ii)[19] *If $\sum_{-\infty}^{\infty} |b_n|^2 < \infty$, there is a function $h(t) \in L^2(-\pi, \pi)$ such that*

$$b_n = \int_{-\pi}^{\pi} e^{i\lambda_n t} h(t)\, dt \qquad (n = 0, \pm1, \pm2, \cdots). \tag{5.3}$$

(iii)[19] *If*

$$b_n = \int_{-\pi}^{\pi} e^{i\lambda_n t} h(t)\, dt \qquad (n = 0, \pm1, \pm2, \cdots), \tag{5.4}$$

with $h(t) \in L^2(-\pi, \pi)$, then

$$\sum_{n=-\infty}^{\infty} |b_n|^2 \leqq A(L) \int_{-\pi}^{\pi} |h(t)|^2 dt.$$

We apply Lemma 8 (ii) to show that a function of exponential type less than π, bounded at points λ_n, is as nearly as possible a finite trigonometric integral.

Lemma 9. *If $f(z)$ is an entire function of exponential type $k < \pi$, and if the numbers $f(\lambda_n)$ are bounded, where λ_n is real, $\lambda_0 = 0$, $|\lambda_n - n| \leqq L < 1/\pi^2$, then*

$$f(z) = f(0) + z \int_{-k}^{k} e^{izt} h(t)\, dt, \qquad h(t) \in L^2(-k, k). \tag{5.5}$$

[18] Paley and Wiener [13], pp. 108, 100.
[19] Boas [3].

We define numbers b_n by

$$b_0 = f'(0); \qquad b_n = \frac{f(\lambda_n) - f(0)}{\lambda_n} \qquad (n = \pm 1, \pm 2, \cdots).$$

Let $h(t)$ be the function, whose existence is assured by Lemma 8 (ii), such that

$$b_n = \int_{-\pi}^{\pi} e^{i\lambda_n t} h(t)\, dt, \qquad h(t) \in L^2(-\pi, \pi),$$

and let

$$H(z) = \int_{-\pi}^{\pi} e^{izt} h(t)\, dt. \tag{5.6}$$

Then, for a suitable number c, the entire function

$$F(z) = \frac{1}{z}\left\{H(z) - \frac{f(z) - f(0)}{z} - c\psi(z)\right\}$$

has a zero at each λ_n ($\psi(z)$ is the function defined by (3.1)). By Lemma 3, we have, for some C,

$$| F(re^{i\theta}) | \leqq Ce^{\pi r}(1 + r)^{-2(1-1/\pi^2)}.$$

Since $2(1 - \pi^{-2}) > \pi^{-2}$, $F(z)$ satisfies the hypotheses of the corollary of Lemma 7, and hence vanishes identically, so that

$$H(z) - \frac{1}{z}\{f(z) - f(0)\} = c\psi(z). \tag{5.7}$$

But c must be zero, since

$$| H(iy) | \leqq \int_{-\pi}^{\pi} e^{-\pi y t} | h(t) |\, dt \leqq \left(\int_{-\pi}^{\pi} e^{-2\pi y t}\, dt\right)^{\frac{1}{2}} \left(\int_{-\pi}^{\pi} | h(t) |^2\, dt\right)^{\frac{1}{2}}$$

$$= O(| y |^{-\frac{1}{2}} e^{\pi | y |}), \qquad | y | \to \infty;$$

while, by Lemma 5,

$$\psi(iy) = \Omega(| y |^{-2/\pi^2} e^{\pi | y |}), \qquad | y | \to \infty,$$

and $2\pi^{-2} < \frac{1}{2}$.

Since $c = 0$, (5.5) follows from (5.7) and (5.6).

6. We can now prove

THEOREM 4. *Let the numbers λ_n be real, $\lambda_0 = 0$, and*

$$| \lambda_n - n | \leqq L < \frac{1}{\pi^2} \qquad (n = \pm 1, \pm 2, \cdots). \tag{6.1}$$

For each non-negative integer r there exist functions $c_n(z)$ and $R_j(z)$ such that if $f(z)$ is an entire function of exponential type $k < \pi$ and the numbers $f(\lambda_n)$ are bounded, then[20]

$$f(z) = \sum_{n=-\infty}^{\infty}{}' c_n(z) f(\lambda_n) + \sum_{j=0}^{r+1} R_j(z) f^{(j)}(0), \tag{6.2}$$

$$\sum_{n=-\infty}^{\infty}{}' |n^{r+1} c_n(z)|^2 \leqq \frac{A(k, r)}{(1 - \pi L^{\frac{1}{2}})^2} e^{2\pi|y|}(1 + |z|)^{2r+2}. \tag{6.3}$$

The functions $R_j(z)$ are independent of $f(z)$ (but depend on the λ_n).

Suppose first that $f(z)$ is of the form

$$f(z) = \int_{-k}^{k} e^{izt}\, d\alpha(t), \qquad \int_{-k}^{k} |d\alpha(t)| < \infty. \tag{6.4}$$

Let $F(z, t) = e^{izt} g(t)$, where $g(t) = 1$ on $(-k, k)$, $g^{(r)}(t)$ is the integral of a function of $L^2(-\pi, \pi)$, and $g^{(s)}(\pi) = g^{(s)}(-\pi) = 0$ for $s = 1, 2, \cdots, r + 1$. Then

$$\frac{\partial^{r+1}}{\partial t^{r+1}} F(z, t) = e^{izt} \sum_{j=0}^{r+1} \binom{r+1}{j} (iz)^j g^{(r-j+1)}(t),$$

and belongs to $L^2(-\pi, \pi)$. By Lemma 8 (i),

$$\frac{\partial^{r+1}}{\partial t^{r+1}} F(z, t) = \operatorname*{l.i.m.}_{N \to \infty} \sum_{n=-N}^{N} b_n(z) e^{i\lambda_n t},$$

with

$$\begin{aligned} \sum_{n=-\infty}^{\infty} |b_n(z)|^2 &\leqq B(L) \int_{-\pi}^{\pi} \left| \frac{\partial^{r+1}}{\partial t^{r+1}} F(z, t) \right|^2 dt \\ &\leqq \frac{A(k, r)}{(1 - \pi L^{\frac{1}{2}})^2} e^{2\pi|y|}(1 + |z|)^{2r+2}. \end{aligned} \tag{6.5}$$

Hence

$$\begin{aligned} F(z, t) &= \sum_{n=-\infty}^{\infty} \frac{b_n(z)}{r!} \int_{-\pi}^{t} (t - u)^r e^{i\lambda_n u}\, du \\ &= \frac{(t + \pi)^{r+1}}{(r + 1)!} b_0(z) + \sum_{n=-\infty}^{\infty}{}' \frac{b_n(z) e^{i\lambda_n t}}{(i\lambda_n)^{r+1}} + \sum_{n=-\infty}^{\infty}{}' b_n(z) e^{-i\lambda_n \pi} \sum_{j=1}^{r} \frac{A_j t^j}{\lambda_n^j}. \end{aligned} \tag{6.6}$$

By (6.5) and Cauchy's inequality, the series on the right converge absolutely and uniformly (for fixed z). We may therefore replace $F(z, t)$ by the right side of (6.6), and integrate term by term with respect to $\alpha(t)$. Since $f(z) = \int_{-\pi}^{\pi} F(z, t)\, d\alpha(t)$, we obtain (6.2) with $c_n(z) = b_n(z)(i\lambda_n)^{-r-1}$; and (6.3) follows from (6.5).

[20] $\sum'$ omits the term for which $n = 0$. The $c_n(z)$ and $R_j(z)$ depend, of course, on r.

We now consider the general case. By Lemma 9, $f(z)$ has the representation (5.5). Take functions $h_\nu(t)$, each of bounded variation on $(-k, k)$, such that $h_\nu(k) = h_\nu(-k) = 0$, and $h(t) = \underset{\nu\to\infty}{\text{l.i.m.}}\, h_\nu(t)$. If we set

$$f_\nu(z) = f(0) + z \int_{-k}^{k} e^{izt} h_\nu(t)\, dt,$$

then $f(z) - f_\nu(z) \to 0$ and $f^{(j)}(0) - f_\nu^{(j)}(0) \to 0$ as $\nu \to \infty$, for $j = 1, 2, \cdots$. Now

$$f_\nu(z) = f(0) + i \int_{-k}^{k} e^{izt}\, dh_\nu(t) = \int_{-k}^{k} e^{izt}\, dh_\nu^*(t),$$

$$h_\nu^*(t) = f(0) \operatorname{sgn} t + ih_\nu(t).$$

Consequently, as we have just shown, $f_\nu(z)$ satisfies (6.2). Since the numbers $f(\lambda_n)$ are bounded, we can form

$$g(z) = \sum_{n=-\infty}^{\infty} c_n(z) f(\lambda_n) + \sum_{j=0}^{r+1} R_j(z) f^{(j)}(0).$$

Then

$$\begin{aligned} |g(z) - f_\nu(z)| \leq \sum_{j=0}^{r+1} |R_j(z)|\,|f^{(j)}(0) - f_\nu^{(j)}(0)| \\ + \sum_{|n|>N} |c_n(z) f(\lambda_n)| + \sum_{|n|>N} |c_n(z) f_\nu(\lambda_n)| \\ + \sum_{|n|\leq N} |c_n(z)|\,|f(\lambda_n) - f_\nu(\lambda_n)| \\ = S_0 + S_1 + S_2 + S_3. \end{aligned}$$

For fixed z, $S_1 \to 0$ as $N \to \infty$. By Lemma 8 (iii),

$$\begin{aligned} S_2^2 &\leq \sum_{|n|>N} |\lambda_n c_n(z)|^2 \sum_{|n|>N} \left| \frac{f_\nu(\lambda_n)}{\lambda_n} \right|^2 \\ &\leq A(L) \sum_{|n|>N} |\lambda_n c_n(z)|^2 \int_{-k}^{k} |h_\nu(t)|^2\, dt. \end{aligned}$$

Since $h(t) = \text{l.i.m.}\, h_\nu(t)$, there is a number B such that

$$\int_{-k}^{k} |h_\nu(t)|^2\, dt \leq B \int_{-k}^{k} |h(t)|^2\, dt \qquad (\nu = 1, 2, \cdots);\text{[21]}$$

hence, by (6.3), $S_2 \to 0$ as $N \to \infty$, uniformly with respect to ν. For fixed N, $S_0 \to 0$ and $S_3 \to 0$ as $\nu \to \infty$. Hence

$$g(z) = \lim_{\nu\to\infty} f_\nu(z) = f(z);$$

that is, (6.2) holds.

[21] Unless $h(t) = 0$ almost everywhere, when there is nothing to prove.

7. We can now establish our generalization of Miss Cartwright's theorem.

THEOREM 5. *If $f(z)$ is an entire function of exponential type $k < \pi$, and if*

$$|f(\lambda_n)| \leqq K \qquad (n = 0, \pm 1, \pm 2, \cdots), \tag{7.1}$$

where the λ_n are real and $|\lambda_n - n| \leqq L < 1/(2\pi^2)$, then

$$|f(x)| \leqq \frac{A(k)K}{1 - \pi(2L)^{\frac{1}{2}}} \qquad (-\infty < x < \infty). \tag{7.2}$$

For any integer m, set $\lambda_n^{(m)} = \lambda_{n+m} - \lambda_m$. Then

$$|\lambda_n^{(m)} - n| = |\lambda_{n+m} - (n + m) - (\lambda_m - m)| \leqq 2L < \pi^{-2}, \qquad \lambda_0^{(m)} = 0.$$

We take the $\lambda_n^{(m)}$ as the λ_n of Theorem 4, and apply Theorem 4, with $r = 0$, to[22]

$$f_m(z) = f(z + \lambda_m) \sin^2 \epsilon z, \qquad \epsilon = \tfrac{1}{4}(\pi - k).$$

We obtain, using Cauchy's inequality and (6.5),

$$|f(x \pm i + \lambda_m) \sin^2 \{\epsilon(x \pm i)\}| \leqq K \sum_{n=-\infty}^{\infty}{}' |c_n(x \pm i)| \leqq \frac{A(k)K}{1 - \pi(2L)^{\frac{1}{2}}},$$

for $|x| \leqq 2$. Since $|\csc^2 \{\epsilon(x \pm i)\}| \geqq A(\epsilon) > 0$, we have

$$|f(x \pm i + \lambda_m)| \leqq \frac{A(k)K}{1 - \pi(2L)^{\frac{1}{2}}} \qquad (|x| \leqq 2; -\infty < m < \infty);$$

or

$$|f(x \pm i)| \leqq \frac{A(k)K}{1 - \pi(2L)^{\frac{1}{2}}} \qquad (-\infty < x < \infty).$$

By a Phragmén-Lindelöf theorem,[23]

$$|f(x)| \leqq \frac{A(k)K}{1 - \pi(2L)^{\frac{1}{2}}} \qquad (-\infty < x < \infty),$$

and this is the desired result.

8. Theorem 5 implies the following apparently more general theorem.

THEOREM 6. *Let $\varphi(t)$ be a non-negative, non-decreasing, unbounded function in $t \geqq 0$. If $f(z)$ is an entire function of exponential type $k < \pi$, and*

$$\sup_{-\infty < n < \infty} \int_{n-\epsilon}^{n+\delta} \varphi(|f(x)|)\, dx = K < \infty, \tag{8.1}$$

with $\delta \geqq 0$, $\epsilon \geqq 0$, $\delta + \epsilon > 0$, then $f(z)$ is bounded on the real axis.

[22] The device is Pfluger's ([14], p. 313).

[23] Titchmarsh [18], p. 180.

More generally, if $|\lambda_n - n| \leqq L < (2\pi^2)^{-1}$, *and* $\alpha(x)$ *is a non-decreasing function having a point of increase in* $|x| \leqq \delta < (2\pi^2)^{-1} - L$, *then*

$$\sup_{-\infty<n<\infty} \int_{-\delta}^{\delta} \varphi(|f(x + \lambda_n)|)\, d\alpha(x) = K < \infty \tag{8.2}$$

implies

$$\sup_{-\infty<x<\infty} |f(x)| \leqq \frac{A(k)}{1 - \pi(2\delta + 2L)^{\frac{1}{2}}} \varphi^{-1}\left\{\frac{K}{\alpha(\delta) - \alpha(-\delta)}\right\} \tag{8.3}$$

(*where the largest value of* $\varphi^{-1}\{\cdots\}$ *is taken if* φ *does not have a single-valued inverse*).

When $\varphi(t) = t$ and $\alpha(x) = \operatorname{sgn} x$, we have Theorem 5 again.

A special case is that $f(z)$ is bounded on the real axis if it is of (any) exponential type, and if for some $\delta > 0$

$$\sup_{-\infty<X<\infty} \int_X^{X+\delta} \varphi(|f(x)|)\, dx < \infty;$$

and a still more special case is that a function of exponential type is bounded on the real axis if[24]

$$\int_{-\infty}^{\infty} \varphi(|f(x)|)\, dx < \infty.$$

If (8.2) is true, there are necessarily points μ_n such that $|\mu_n - \lambda_n| \leqq \delta$, and so $|\mu_n - n| \leqq \delta + L < (2\pi^2)^{-1}$, and

$$\varphi(|f(\mu_n)|) \leqq \frac{K}{\alpha(\delta) - \alpha(-\delta)}.$$

Then

$$|f(\mu_n)| \leqq \varphi^{-1}\left\{\frac{K}{\alpha(\delta) - \alpha(-\delta)}\right\},$$

and (8.3) follows from Theorem 5.

In particular, if (8.1) is satisfied, and $\delta < \frac{1}{20}$, $\epsilon < \frac{1}{20}$ (which involves no real loss of generality), we have $L = 0$, and

$$\sup_{-\infty<x<\infty} |f(x)| \leqq A(k) \sup_{-\infty<n<\infty} \varphi^{-1}\left\{\frac{1}{\delta + \epsilon} \int_{n-\epsilon}^{n+\delta} \varphi(|f(t)|)\, dt\right\}.$$

We may express this by saying that $f(z)$ is bounded on the real axis if it has an average which is bounded near the integers.

We now establish a generalization of Theorem 3 in which n is replaced by λ_n, and $p > 0$ instead of $p \geqq 1$.

THEOREM 7. *Let* $\beta(x)$ *be a non-decreasing function defined on* $(-\infty, \infty)$, *such that* $\beta(x + 1) - \beta(x) \leqq B < \infty$. *Let* $\gamma(t)$ *be a non-decreasing function*

[24] For $\varphi(t) = t^p$, $p > 0$, this was proved in another way by Plancherel and Pólya ([15], p. 124). See also footnote 5.

defined on $(-\delta, \delta)$, $\delta > 0$. *Let* λ_n *be real, and* $|\lambda_n - n| \leqq L < (2\pi^2)^{-1}$. *If* $f(z)$ *is an entire function of exponential type* $k < \pi$, *and* $p > 0$, *then*

$$\left\{\int_{-\infty}^{\infty} |f(x)|^p \, d\beta(x)\right\}^{1/p} \tag{8.4}$$

$$\leqq \frac{A(k, p)(3 + \delta)^{A(p)} B^{1/p}}{1 - \pi(2L)^{1/2}} \left\{\sum_{n=-\infty}^{\infty} \frac{1}{\gamma(\delta) - \gamma(-\delta)} \int_{-\delta}^{\delta} |f(\lambda_n + t)|^p \, d\gamma(t)\right\}^{1/p}.$$

If $p \geqq 1$, *the functions* $A(k, p)$ *and* $A(p)$ *may be taken to be independent of* p.

The special case $\beta(x) = x$, $\gamma(t) = \operatorname{sgn} t$ (corresponding to (i) of Theorem 3) is particularly interesting. The analogue of (ii) of Theorem 3, however, gives us nothing new.

If the right side of (8.4) converges, $\sum_{\nu=-\infty}^{\infty} |f(\lambda_\nu + t)|^p$ converges for almost all t on $(-\delta, \delta)$ (i.e., almost all with respect to $\gamma(t)$). We confine our attention to values of t for which the series converges.

We write $\lambda_n^{(m)} = \lambda_{n+m} - \lambda_m$ ($m, n = 0, \pm 1, \pm 2, \cdots$), and $f_m(w) = f(w + \lambda_m + t)(\sin \zeta w)^r$, where $r = 2 + [1/p]$, and $\zeta = (\pi - k)/(2r)$. We apply Theorem 4 to $f_m(w)$, taking the $\lambda_n^{(m)}$ for the λ_n of Theorem 4, and then putting $w = z - t$, where $|t| \leqq \delta$, $|x| \leqq 1 + \delta$, $|y| \leqq 1$. If $c_\nu^{(m)}(w)$ is the function $c_\nu(w)$ associated with $\lambda_\nu^{(m)}$, we have, using (6.5),

$$\begin{aligned} |f_m(w)| &\leqq \sum_\nu{}' |c_\nu^{(m)}(w)| \, |f_m(\lambda_\nu^{(m)})| \\ &\leqq \left\{\sum_\nu{}' |c_\nu^{(m)}(w)\nu^r|^2\right\}^{1/2} \left\{\sum_\nu{}' \left|\frac{f_m(\lambda_\nu^{(m)})}{\nu^r}\right|^2\right\}^{1/2} \\ &\leqq \frac{A(k, r)(1 + |w|)^r}{1 - \pi(2L)^{1/2}} \sum_\nu{}' \left|\frac{f_m(\lambda_\nu^{(m)})}{\nu^r}\right|. \end{aligned} \tag{8.5}$$

When z is on the boundaries of the rectangles defined by

$$|y| \leqq 1, \qquad \left|x - t - \frac{n\pi}{\zeta}\right| \leqq \frac{\pi}{2\zeta} \qquad (n = 0, \pm 1, \pm 2, \cdots),$$

we have $|\sin \zeta(z - t)|^{-r} < A(r, k) = A(p, k)$. For such z, we may drop the factor $|\sin \zeta(z - t)|^r$ from the left side of (8.5); the resulting inequality, by the maximum principle, holds inside the rectangles as well. Consequently we have, for $|x| \leqq 1$,

$$|f(x + \lambda_m)| \leqq \frac{A(k, p)(3 + \delta)^r}{1 - \pi(2L)^{1/2}} \sum_\nu{}' \left|\frac{f(\lambda_{\nu+m} + t)}{\nu^r}\right|. \tag{8.6}$$

If $p \geqq 1$ (when $r = 2$), the rest of the proof is almost identical with that of Theorem 3, and is omitted. We now complete the proof in the case $0 < p < 1$.

We take the p-th power of both sides of (8.6), apply Jensen's inequality to the series on the right, and integrate over $(-\delta, \delta)$ with respect to $\gamma(t)$. The result is

$$|f(x + \lambda_m)|^p \leqq A(k, p)C^p \sum_\nu{}' |\nu|^{-(r+1)p} \int_{-\delta}^{\delta} |f(\lambda_{\nu+m} + t)|^p \, d\gamma(t),$$

where

$$C = \frac{(3 + \delta)^{r+1}}{\{1 - \pi(2L)^{\frac{1}{2}}\}\{\gamma(\delta) - \gamma(-\delta)\}}.$$

Then

$$\int_{-\infty}^{\infty} |f(x)|^p \, d\beta(x) \leqq \sum_{m=-\infty}^{\infty} \int_{-1}^{1} |f(x + \lambda_m)|^p \, d\beta(x + \lambda_m)$$

$$\leqq A(k, p)BC^p \sum_{m=-\infty}^{\infty} \sum_\nu{}' \frac{1}{|\nu|^{(r+1)p}} \int_{-\delta}^{\delta} |f(\lambda_{\nu+m} + t)|^p \, d\gamma(t)$$

$$\leqq A(k, p)BC^p \sum_{m=-\infty}^{\infty} \int_{-\delta}^{\delta} |f(\lambda_m + t)|^p \, d\gamma(t).$$

This completes the proof.

9. As an additional application of Lemma 7, we prove a theorem of Levinson's on the completeness of $\{e^{i\lambda_n t}\}$.

THEOREM 8.[25] *If* $|\lambda_n - n| \leqq (p - 1)/(2p)$, *the functions* $\{e^{i\lambda_n t}\}$ $(n = 0, \pm 1, \pm 2, \cdots)$ *form a complete set in* $L^p(-\pi, \pi)$ $(1 < p < \infty)$.[26]

The set $\{e^{i\lambda_n t}\}$ is, by definition, complete in L^p if any $f(t) \epsilon L^p$ such that

$$\text{(9.1)} \qquad \int_{-\pi}^{\pi} e^{i\lambda_n t} f(t) \, dt = 0 \qquad (n = 0, \pm 1, \pm 2, \cdots)$$

is zero almost everywhere. Suppose that (9.1) is satisfied, and set

$$H(z) = \int_{-\pi}^{\pi} e^{izt} f(t) \, dt,$$

so that $H(\lambda_n) = 0$. We have, for $|y| \geqq 1$,

$$|H(x + iy)| \leqq \int_{-\pi}^{\pi} e^{-yt} |f(t)| \, dt$$

$$\leqq \left(\int_{-\pi}^{\pi} |f(t)|^p \, dt\right)^{1/p} \left(\int_{-\pi}^{\pi} e^{-p'yt} \, dt\right)^{1/p'} \qquad \left(p' = \frac{p}{p-1}\right)$$

$$\leqq B|y|^{-1/p'} e^{\pi|y|},$$

[25] Levinson [11] proves a still more precise theorem.

[26] The theorem remains true when $p = \infty$, and the same proof establishes it if expressions involving p are interpreted according to the usual conventions; we must, however, assume not only that $|\lambda_n - n| \leqq \frac{1}{2}$ but also that $\lambda_n \neq \lambda_{n+1}$ $(n = 0, \pm 1, \pm 2, \cdots)$.

with B independent of x and y. Lemma 7 shows that $H(z) \equiv c\psi(z)$, where $\psi(z)$ is defined by (3.1), with the given λ_n's. By Lemma 6, $|\psi(x + 3i)| > A > 0$ $(-\infty < x < \infty)$. Hence, if $c \neq 0$,

$$|H(x + 3i)| > A\,|c| > 0 \qquad (-\infty < x < \infty).$$

But, by the Riemann-Lebesgue theorem,

$$H(x + 3i) = \int_{-\pi}^{\pi} e^{ixt} e^{-3yt} f(t)\, dt \to 0, \qquad |x| \to \infty.$$

Consequently $c = 0$, $H(z) \equiv 0$, and so $f(t) = 0$ almost everywhere.

Another way of expressing Theorem 8 is to say that the set $\{e^{i\lambda_n t}\}$ is closed in L^p if $|\lambda_n - n| \leqq (2p)^{-1}$ $(1 < p < \infty)$. That is, any function of $L^p(-\pi, \pi)$ can be approximated, in the metric of L^p, by linear combinations of elements of the set. This leads one to ask whether the set $\{e^{i\lambda_n t}\}$ possesses a "Weierstrass's theorem": that is, whether every continuous function can be uniformly approximated by linear combinations of the $e^{i\lambda_n t}$. Since, as Levinson has proved [10], the formal development of (in particular) a continuous function in terms of the $e^{i\lambda_n t}$ is equisummable with its Fourier series, if $|\lambda_n - n| \leqq \delta < \frac{1}{4}$, it follows that on any interval $(-\pi + \epsilon, \pi - \epsilon)$, $\epsilon > 0$, a continuous function can be approximated uniformly by linear combinations of the $e^{i\lambda_n t}$ (if $|\lambda_n - n| \leqq \delta < \frac{1}{4}$). However, a little more than this is true. We shall prove

THEOREM 9. *If $|\lambda_n - n| \leqq \frac{1}{4}$ and $\epsilon > 0$, any function continuous in $-\pi \leqq t \leqq \pi - \epsilon$ can be uniformly approximated by linear combinations of the $e^{i\lambda_n t}$.*

This rather curious result of course neither implies nor is implied by Levinson's equiconvergence theorem.

Any function continuous in $(-\pi, \pi - \epsilon)$ can be uniformly approximated by linear combinations of the $e^{i\lambda_n t}$ provided that

$$\int_{-\pi}^{\pi-\epsilon} e^{i\lambda_n t}\, d\alpha(t) = 0 \qquad (n = 0, \pm 1, \pm 2, \cdots), \tag{9.2}$$

with $\alpha(t)$ of bounded variation on $(-\pi, \pi - \epsilon)$, implies that $\alpha(t)$ differs from a constant at most on a countable set.[27]

Suppose that (9.2) is satisfied, and set

$$H(z) = \int_{-\pi}^{\pi-\epsilon} e^{izt}\, d\alpha(t).$$

Then $H(z)$ is an entire function vanishing at $z = \lambda_n$, and

$$|H(x + iy)| \leqq e^{\pi|y|} \int_{-\pi}^{\pi-\epsilon} |d\alpha(t)| \leqq Ce^{\pi|y|},$$

say. Consider the entire function

$$G(z) = \frac{H(z) - h\psi(z)}{z - \lambda_0},$$

[27] See Banach [1], p. 73, where the result is stated in a different (but equivalent) form.

where $h = H'(\lambda_0)/\psi'(\lambda_0)$ is chosen so that $G(\lambda_0) = 0$. Then $G(\lambda_n) = 0$ for $n = 0, \pm 1, \pm 2, \cdots$, and

$$| G(x + iy) | \leqq \frac{C + A | h | (1 + | y |)^{2\delta}}{| z - \lambda_0 |} e^{\pi |y|}$$
$$= O(| y |^{-\frac{1}{2}} e^{\pi|y|}), \qquad | y | \to \infty.$$

Consequently, by Lemma 7, $G(z) = B\psi(z)$ for some B, and

$$(9.3) \qquad H(z) = (B_1 z + B_2)\psi(z)$$

for some B_1 and B_2.

Now if $y < 0$,

$$(9.4) \qquad | H(iy) | \leqq \int_{-\pi}^{\pi-\epsilon} e^{-yt} | d\alpha(t) | \leqq e^{(\pi-\epsilon)|y|} \int_{-\pi}^{\pi-\epsilon} | d\alpha(t) |.$$

But by Lemma 5, $\psi(iy) = \Omega(| y |^{-\frac{1}{2}} e^{\pi|y|})$, $| y | \to \infty$, and so $H(iy) = \Omega(| y |^{-\frac{1}{2}} e^{\pi|y|})$, by (9.3), unless $B_1 = B_2 = 0$. But (9.4) shows that $H(iy) = o(| y |^{-\frac{1}{2}} e^{\pi|y|})$ as $y \to -\infty$; therefore $B_1 = B_2 = 0$, and $H(z) \equiv 0$. Consequently $\alpha(t)$ is constant except at most on a countable set,[28] and the theorem is proved.

References

1. S. Banach, *Théorie des Opérations Linéaires*, 1932.
2. R. P. Boas, Jr., *Representations for entire functions of exponential type*, Annals of Mathematics, (2), vol. 39(1938), pp. 269–286.
3. R. P. Boas, Jr., *A trigonometric moment problem*, Journal of the London Mathematical Society, vol. 14(1939), pp. 242–244.
4. S. Bochner, *Vorlesungen über Fouriersche Integrale*, 1932.
5. M. L. Cartwright, *On certain integral functions of order one*, Quarterly Journal of Mathematics (Oxford series), vol. 7(1936), pp. 46–55.
6. A. Gelfond, *Interpolation et unicité des fonctions entières*, Matematicheskii Sbornik (Recueil Mathématique), new series, vol. 4(1938), pp. 115–147.
7. G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, 1934.
8. V. G. Iyer, *On the Lebesgue class of integral functions along straight lines issuing from the origin*, Quarterly Journal of Mathematics (Oxford series), vol. 7(1936), pp. 294–299.
9. N. Levinson, *On a problem of Pólya*, American Journal of Mathematics, vol. 58(1936), pp. 791–798.
10. N. Levinson, *On non-harmonic Fourier series*, Annals of Mathematics, (2), vol. 37(1936), pp. 919–936.
11. N. Levinson, *On the closure of* $\{e^{i\lambda_n x}\}$, this Journal, vol. 2(1936), pp. 511–516.
12. A. J. Macintyre, *Laplace's transformation and integral functions*, Proceedings of the London Mathematical Society, (2), vol. 45(1938–39), pp. 1–20.
13. R. E. A. C. Paley and N. Wiener, *Fourier Transforms in the Complex Domain*, 1934.
14. A. Pfluger, *On analytic functions bounded at the lattice points*, Proceedings of the London Mathematical Society, (2), vol. 42(1936–37), pp. 305–315.
15. M. Plancherel and G. Pólya, *Fonctions entières et intégrales de Fourier multiples*, Commentarii Mathematici Helvetici, vol. 9(1936–37), pp. 224–248; vol. 10(1937–38), pp. 110–163.

[28] See Bochner [4], p. 67.

16. G. PÓLYA AND G. SZEGÖ, *Aufgaben und Lehrsätze aus der Analysis*, 1925.
17. C. L. SIEGEL, *Über Gitterpunkte in convexen Körpern und ein damit zusammenhängendes Extremalproblem*, Acta Mathematica, vol. 65(1935), pp. 307–323.
18. E. C. TITCHMARSH, *The Theory of Functions*, 1932.
19. G. VALIRON, *Sur la formule d'interpolation de Lagrange*, Bulletin des Sciences Mathématiques, (2), vol. 49(1925), pp. 181–192, 203–224.
20. J. M. WHITTAKER, *Interpolatory Function Theory*, 1935.

CAMBRIDGE, ENGLAND.

FUNCTIONS OF SEVERAL VARIABLES AND ABSOLUTE CONTINUITY, I

By J. W. Calkin

Introduction. We are concerned in this paper with certain properties of real- and complex-valued functions of n real variables which are "potential functions of their generalized derivatives"—a concept introduced by G. C. Evans [2, 3][1] for functions of two variables and readily extended to the case of n variables, $n \geqq 2$. We shall call these functions merely functions of class $\mathfrak{P}$, sacrificing in the interests of brevity the descriptiveness of Evans' terminology.

While the results which are described here are not without intrinsic interest, their chief importance for us lies in the uses to which they are put in researches to be described in subsequent papers: the author of the present paper has found them necessary for the study of partial differential equations by means of the theory of transformations in Hilbert space and the author of the following paper, *Functions of several variables and absolute continuity*, II (C. B. Morrey, Jr.) has found them necessary for certain investigations in the calculus of variations. From the point of view of both of these branches of research, the importance of functions of class $\mathfrak{P}$ lies in the fact that they arise in the following way: Let $\{f_k(x_1, \cdots, x_n\}$ be a sequence of functions (real- or complex-valued) of class C′ on an open set G in n-dimensional space, and let the sequences $\{f_k\}$, $\{\partial f_k/\partial x_j\}$ $(j = 1, 2, \cdots, n)$ converge in the mean of order p $(p \geqq 1)$ on G. Then the limit (in the mean) f of the sequence $\{f_k\}$ is of class $\mathfrak{P}$ and the limits of the sequences of partial derivatives are the generalized derivatives of f in the sense of Evans. This result gives rise to several interesting questions. Most important, while the function f described above can evidently be arbitrarily defined on any set of measure zero, and thus cannot be expected to have partial derivatives in the ordinary sense, it is natural to seek to use this very freedom to redefine the function on a set of measure zero so as to obtain a function in some sense differentiable. It is shown below that the problem which arises in this connection has an entirely satisfactory solution; any function of class $\mathfrak{P}$ is equal almost everywhere to a function f_0 which has the following property: on almost all lines parallel to the x_j-axis and intersecting G, f_0 is absolutely continuous on every closed interval interior to G $(j = 1, \cdots, n)$. Functions of class $\mathfrak{P}$ which have this property are introduced below as functions of class $\mathfrak{P}'$.

We may also mention here that the convergence theorem stated above remains valid if the functions f_k are required merely to be of class $\mathfrak{P}$, and the sequences of derivatives are replaced by sequences of generalized derivatives. This im-

Received September 12, 1939.

[1] Numbers in brackets refer to the bibliography at the end of the paper.

portant "closure" property of the class $\mathfrak{P}$ together with related "compactness" properties is fundamental for the application of certain topological methods in the calculus of variations for multiple integrals and for the use of the theory of Hilbert space in the study of elliptic partial differential equations. Furthermore, coupled with the result mentioned concerning the differentiability properties of functions of class $\mathfrak{P}$, it suggests that the class $\mathfrak{P}'$ may, in many branches of analysis, properly be regarded as the analogue, for functions of n variables, of the class of absolutely continuous functions of a single variable. It is to be emphasized, however, that a function of class $\mathfrak{P}'$ is in general neither absolutely continuous in the sense of Tonelli nor equivalent to such a function.

Apart from the introduction, the present paper is divided into five sections. In the first, various notational and terminological conventions are introduced, and in the second a few facts are summarized which play an important auxiliary rôle, but which lie outside the central train of thought of the paper. In §3, we prepare the way for the introduction of functions of class $\mathfrak{P}$ which are characterized by identical conditions with respect to each of the coördinate variables, by studying in detail the implications of this condition for a single coördinate variable. In §4, functions of class $\mathfrak{P}$ and $\mathfrak{P}'$ are introduced and the relation of this notion to absolute continuity in the sense of Tonelli is discussed along lines essentially laid down by Evans. In §5 the imposition of certain integrability conditions on functions of class $\mathfrak{P}$ is considered and a variant of the closure theorem mentioned above is proved.

1. **Notation and terminology.** In order to keep our exposition reasonably concise, we now introduce certain conventions for use in the sequel.

In the first place, we shall be dealing throughout the paper with functions defined on an open set in n-dimensional Euclidean space; accordingly the letter G always denotes an open set which may be either bounded or not, and may, in particular, be the entire space.

Secondly, whenever possible we use the single letter x to denote the Cartesian coördinates $(x_1, x_2, \cdots, x_n)$ of a point in n-dimensional space. In accordance with this convention, the n-dimensional closed *cell*, $a_1 \leqq x_1 \leqq b_1$, $a_2 \leqq x_2 \leqq b_2, \cdots, a_n \leqq x_n \leqq b_n$ is denoted by $[a, b]$ and the integral of a function $f(x_1, x_2, \cdots, x_n)$ over $[a, b]$ by $\int_a^b f\,dx$. Thus, in general, we write merely $f(x)$ for a function $f(x_1, x_2, \cdots, x_n)$ defined on G. However, since we sometimes find it necessary to consider a single coördinate variable x_k, or one of the coördinate hyperplanes, we supplement this convention with further ones as follows: If $x = (x_1, x_2, \cdots, x_n)$ is a point in n-space, we denote by x_k' the set of coördinates $(x_1, x_2, \cdots, x_{k-1}, x_{k+1}, \cdots, x_n)$, the value of x_k being fixed or variable as the purposes of our argument require. Thus if $[a, b]$ has the meaning described above, $[a_k', b_k']$ denotes the $(n-1)$-dimensional cell $a_1 \leqq x_1 \leqq b_1, \cdots, a_{k-1} \leqq x_{k-1} \leqq b_{k-1}, a_{k+1} \leqq x_{k+1} \leqq b_{k+1}, \cdots, a_n \leqq x_n \leqq b_n$, and $\int_{a_k'}^{b_k'} f\,dx_k'$

denotes the integral of f over $[a_k', b_k']$. Furthermore, if it is desired to consider the behavior of a function $f(x)$ with reference to a particular coördinate x_k, or on the hyperplanes orthogonal to the x_k-axis, we write (x_k', x_k) for x and $f(x_k', x_k)$ for $f(x)$. In particular, if the coördinates x_k' are fixed $x_k' = a_k'$, then $f(a_k', x_k)$ denotes the function $f(a_1, \cdots, a_{k-1}, x_k, a_{k+1}, \cdots, a_n)$ of the single variable x_k; similarly, if x_k is fixed, $x_k = a_k$, $f(x_k', a_k)$ denotes a function of the $n - 1$ variables x_k'.

In the preceding discussion and throughout the paper, except where the two cases are explicitly differentiated, the functions $f(x)$ considered may be taken as either real- or complex-valued.

Finally, we point out that we use the words *measurable, integrable,* and related terms, to refer to ordinary Lebesgue integration.

2. **Average functions.** If $f(x)$ is measurable on the open set G, and if h is a small positive constant, the function

$$(2h)^{-n} \int_{x_1-h}^{x_1+h} \cdots \int_{x_n-h}^{x_n+h} f(\xi_1, \cdots, \xi_n)\, d\xi_1 \cdots d\xi_n,$$

defined at every point x of G for which the integral exists, provides an extremely useful approximation to $f(x)$. We call this function the h-average of $f(x)$ and denote it by $f_h(x)$. For brevity, we write $f_h(x)$ in the form

$$(2h)^{-n} \int_{x-h}^{x+h} f(\xi)\, d\xi;$$

in further accord with the conventions previously introduced, we replace this integral by

$$\int_{x_k'-h}^{x_k'+h} \int_{x_k-h}^{x_k+h} f(\xi_k', \xi_k)\, d\xi_k'\, d\xi_k$$

or by the similar iterated integral with the order of integration reversed, whenever the occasion makes it convenient.

The following theorems concerning h-averages find frequent application in the following pages.

THEOREM 2.1. *Let $f(x)$ be measurable on G and integrable on every closed cell interior to G. Then the function $f_h(x)$ is continuous in x on the set where it is defined.*

THEOREM 2.2. *Let $f(x)$ be as in Theorem 2.1. Then almost everywhere on G, $\lim_{h \to 0} f_h(x)$ exists and is equal to $f(x)$.*

THEOREM 2.3. *Let $f(x)$ be measurable on G and let the integral $\int_G |f(x)|^p\, dx$ exist for some $p \geqq 1$. Let G_h denote the set of points x of G such that the closed cell $[x - h, x + h]$ is interior to G. Then f_h is defined on G_h and*

$$\int_{G_h} |f_h(x)|^p\, dx \leqq \int_G |f(x)|^p\, dx. \tag{2.1}$$

If F is an arbitrary bounded closed set interior to G, there exists a positive constant h_0 such that $F \subset G_h$ for $h \leqq h_0$, and

$$\lim_{h\to 0} \int_F |f(x) - f_h(x)|^p \, dx = 0. \tag{2.2}$$

Theorem 2.1 is obvious and Theorem 2.2 follows at once from the familiar theorem of Lebesgue on absolutely continuous set-functions [8]. Theorem 2.3 is known; C. B. Morrey has given a proof of essentially the same proposition for the case of a real-valued function of two variables [5]. A proof for the more general case considered here can be modeled on the latter.

For further properties of the h-average of an integrable function, we refer the reader to a paper of H. E. Bray [1].

3. **Essentially absolutely continuous functions.** For convenience we now introduce the formal definition:

Definition 3.1. Let G be an open set in n-space and let $\tilde{G}$ be the $2n$-dimensional set of points (c, d), $c_i < d_i$ $(i = 1, 2, \cdots, n)$ such that the closed cell $[c, d]$ belongs to G. A class $\mathcal{C}$ of cells $[a, b]$ is said to constitute almost all cells of G if the set of points (a, b) of $\tilde{G}$ for which $[a, b]$ is not in $\mathcal{C}$ has measure zero in $\tilde{G}$.

We can now introduce our first fundamental concept.

Definition 3.2. A function $f(x)$ defined on an open set G is said to be *essentially absolutely continuous* (abbreviated *E.A.C.*) *in* x_k on G if the following conditions are satisfied:

(1) $f(x)$ is measurable on G and integrable on each closed cell interior to G;

(2) there exists a function $g_k(x)$ satisfying the condition (1) above such that the equation

$$\int_{a_k'}^{b_k'} [f(x_k', b_k) - f(x_k', a_k)] \, dx_k' = \int_a^b g_k(x) \, dx \tag{3.1}$$

holds for almost all cells $[a, b]$ of G.

If $n = 1$, we interpret the left member of equation (3.1) to mean $f(b_1) - f(a_1)$. Thus if $n = 1$, it is not difficult to see that $f(x)$ is equal almost everywhere to an absolutely continuous function and, indeed, the obvious generalization of this statement is true for arbitrary n. In order to state this generalization precisely, it is convenient to introduce the following definition:

Definition 3.3. A function f, which on every closed cell $[a, b]$ interior to G is absolutely continuous in x_k for almost all fixed values of x_k' in $[a_k', b_k']$ and which is integrable on every such cell together with $\partial f/\partial x_k$, is said to be *linearly absolutely continuous* (*L.A.C.*) *in* x_k on G.

We shall prove later that any function f which is E.A.C. in x_k on G is equal almost everywhere on G to a function which is L.A.C. in x_k on G. For the present, we confine ourselves to the following result:

THEOREM 3.1. *Let f be E.A.C. in x_k on G and let $[A, B]$ be any cell interior to G. Then there exists a function f_0 which coincides with f almost everywhere on $[A, B]$ and is L.A.C. in x_k on $[A, B]$. Moreover, if $g_k(x)$ is the function associated with f by Definition 3.2, $\partial f_0/\partial x_k = g_k(x)$ almost everywhere on $[A, B]$. If f is L.A.C. on G, it is also E.A.C. on G.*

We first recall that, by Definition 3.2, formula (3.1) holds for all cells $[a, b]$ in $[A, B]$ such that the point (a, b) $(a_i < b_i)$ does not lie in a set E of $2n$-dimensional measure zero. Thus if (a_k, b_k) is not in a certain set E_2' of 2-dimensional measure zero (with $A_k \leqq a_k < b_k \leqq B_k$), we conclude from Fubini's theorem that the set $E(a_k, b_k)$ of points (a_k', b_k'), such that (a, b) is in E, is of $(2n - 2)$-dimensional measure zero. Let Z be the set of values of x_k for which $f(x_k', x_k)$ is not integrable in x_k' over $[A_k', B_k']$; Z is of linear measure zero by Fubini's theorem. Let E_2 be the set E_2' plus the set of points (a_k, b_k) such that either a_k or b_k belongs to Z; then E_2 is still of 2-dimensional measure zero. Moreover, if (a_k, b_k) is not in E_2, the integrals on both sides of (3.1) are continuous in (a_k', b_k') on $[A_k', B_k']$. Thus we conclude that (3.1) holds for every cell $[a, b]$ in $[A, B]$ for which (a_k, b_k) is not in E_2.

Now, let $\bar{a}_k$ be a number between A_k and B_k such that the set of values of b_k for which $(\bar{a}_k, b_k)$ is in E_2 is of linear measure zero. Such a choice is possible by Fubini's theorem. Clearly $\bar{a}_k$ is not in Z. We now define

$$f_0(x_k', x_k) = f(x_k', \bar{a}_k) + \int_{\bar{a}_k}^{x_k} g_k(x_k', \xi_k)\, d\xi_k$$

for each x_k' for which $g_k(x_k', x_k)$ is integrable in x_k (this being almost all x_k'); and define $f_0 = 0$ elsewhere. Clearly f_0 is measurable on $[A, B]$ and A.C. in x_k for almost all x_k' on $[A_k', B_k']$. Also $\partial f_0/\partial x_k$ exists and is equal to $g_k(x)$ almost everywhere and is therefore integrable. By integrating the right side of the above equation with respect to x_k' and using Fubini's theorem, we conclude that $f_0(x_k', x_k)$ is integrable with respect to x_k' for each x_k and that

$$\int_{A_k'}^{B_k'} | f_0(x_k', x_k) |\, dx_k' \leqq \int_{A_k'}^{B_k'} | f(x_k', \bar{a}_k) |\, dx_k' + \int_A^B | g_k(x) |\, dx$$

independently of x_k. Thus it follows that f_0 is integrable over $[A, B]$ and hence that it is L.A.C. on the interior of $[A, B]$. Also, formula (3.1), with f replaced by f_0, holds for every cell $[a, b]$ in $[A, B]$. Thus if b_k is not in a set of measure zero, $f_0(x_k', b_k)$ and $f(x_k', b_k)$ are both integrable in x_k' and

$$\int_{a_k'}^{b_k'} f(x_k', b_k)\, dx_k' = \int_{a_k'}^{b_k'} f_0(x_k', b_k)\, dx_k'$$

for every cell $[a_k', b_k']$ in $[A_k', B_k']$. Therefore $f(x_k', b_k) = f_0(x_k', b_k)$ for almost every x_k' in $[A_k', B_k']$. Hence, as f and f_0 are both integrable on $[A, B]$, it follows that they coincide on $[A, B]$ except on a set of measure zero.

The final statement of the theorem is obvious.

COROLLARY. *If f is E.A.C. in x_k on G and I denotes the set of values of x_k such that, for some x_k', $(x_k'; x_k)$ is in G, then except possibly for a_k or b_k in a set of measure zero in I, equation* (3.1) *is valid for all closed cells* $[a, b]$ *interior to G.*

The straightforward proof of the corollary is left to the reader.

It is of course evident that the requirements of Definition 3.1 could be strengthened so as to render Theorem 3.1 and the above corollary almost obvious; and for the purposes immediately at hand this procedure would be entirely satisfactory. However, in several applications of the present work to be described in later publications of Morrey, functions occur which clearly satisfy the conditions of Definition 3.1, while it is not at all obvious that they satisfy any stronger ones.

We may point out in connection with Theorem 3.1 that it is one of the objects of the present section to establish that a function which is E.A.C. in each of several variables is equal almost everywhere to a single function which is L.A.C. on G with respect to each of the variables in question.

Before turning to that aspect of our discussion, we observe that a "derivative with respect to x_k" can be assigned to a function f which is E.A.C. in x_k, without taking account of the function f_0 described in Theorem 3.1. Here our procedure follows closely that of Evans in a similar discussion ([2], pp. 275–277).

We first note that, by virtue of the corollary to Theorem 3.1, f determines uniquely the function $\int_a^b g_k \, dx$ of closed cells $[a, b]$ interior to G. Thus, by extension, f determines uniquely the absolutely continuous function

$$I(E; f, k) = \int_E g_k \, dx \tag{3.2}$$

of sets E, this function being defined at least for all bounded measurable sets E interior to G and at positive distance from its complement. Furthermore this absolutely continuous set-function has a derivative—in the sense of a well-known definition of Lebesgue (see, for example, [8])—which is a point function defined almost everywhere on G and equal almost everywhere to g_k. Hence, since this derivative is uniquely determined by f, we can introduce

DEFINITION 3.4. Let f be E.A.C. in x_k and let g_k be the function associated with f by Definition 3.2, (2). Then the derivative of the set-function $I(E; f, k)$ defined by equation (3.2) is called the *generalized derivative of f with respect to* x_k and is denoted by $D_{x_k} f$.

In view of the argument preceding Definition 3.4, we can also state

THEOREM 3.2. *Let f be a function E.A.C. in x_k. Then any function g_k which with reference to f satisfies the final condition of Definition* 3.2, *satisfies the equation $g_k = D_{x_k} f$ almost everywhere on G.*

We now proceed to an examination of the properties of the average functions f_h of a function absolutely continuous in one of the coördinate variables. The

two simple lemmas which we prove concerning the average functions have several important implications, not only for functions E.A.C. in one coördinate variable, but for functions which have that property with respect to several variables and thus, in particular, for "potential functions of their generalized derivatives".

LEMMA 3.1. *Let $f(x)$ be E.A.C. in x_k on G and let G_h have the same meaning as in Theorem 2.3. Then the h-average function f_h of f is continuous on G_h, $\partial f_h/\partial x_k$ exists and is continuous at every point of G_h and is given there by the formula*

$$\frac{\partial f_h}{\partial x_k} = (2h)^{-n} \int_{x-h}^{x+h} D_{x_k} f \, dx. \tag{3.3}$$

That f_h is continuous on G_h has already been noted in Theorem 2.1. Furthermore, by the same theorem, the right member of formula (3.2) is continuous. Hence, to complete the proof, we begin by forming the difference quotient

$$\frac{\Delta f_h}{\alpha} = \frac{f_h(a_k', a_k + \alpha) - f_h(a_k', a_k)}{\alpha} = \alpha^{-1}(2h)^{-n} \int_{a-h}^{a+h} [f(x_k', x_k + \alpha) - f(x_k', x_k)] \, dx$$

at an arbitrary point a of G_h. Then, making use of the corollary to Theorem 3.1, we have

$$\frac{\Delta f_h}{\alpha} = \alpha^{-1}(2h)^{-n} \int_{a_k-h}^{a_k+h} \left[\int_{x_k}^{x_k+\alpha} \int_{a_k'-h}^{a_k'+h} D_{x_k} f(x_k', \xi_k) \, dx_k' \, d\xi_k \right] dx_k.$$

By introducing a change in the variable of integration and then employing the theorem of Fubini to change the order of integration, we obtain from the preceding equation

$$\frac{\Delta f_h}{\alpha} = \alpha^{-1}(2h)^{-n} \int_0^{\alpha} \left[\int_{a_k-h}^{a_k+h} \int_{a_k'-h}^{a_k'+h} D_{x_k} f(x_k', x_k + \eta_k) \, dx_k' \, dx_k \right] d\eta_k.$$

Finally, setting $s_k = x_k + \eta_k$, we have

$$\frac{\Delta f_h}{\alpha} = \alpha^{-1} \int_0^{\alpha} (2h)^{-n} \left[\int_{a_k-h+\eta_k}^{a_k+h+\eta_k} \int_{a_k'-h}^{a_k'+h} D_{x_k} f(x_k', s_k) \, dx_k', \, ds_k \right] d\eta_k$$

and it is now necessary only to allow α to approach zero as limit to obtain the formula of the lemma, the existence of $\partial f_h/\partial x_k$ being established in the process.

LEMMA 3.2. *Let $f(x)$ be E.A.C. in x_k on G. Then, if f_h denotes the h-average of f,*

$$f_0(x) = \lim_{h \to 0} f_h(x)$$

is L.A.C. in x_k on G and

$$\frac{\partial f_0}{\partial x_k} = D_{x_k} f$$

almost everywhere on G.

We know already, from Theorem 2.2, that $f_h(x)$ converges almost everywhere to $f(x)$. From Theorem 3.1, we know that if $[A, B]$ is any cell interior to G, $f(x)$ is equal almost everywhere on $[A, B]$ to a function which is L.A.C. on the interior of $[A, B]$. Thus we know that $f_h(x)$ converges almost everywhere on each such cell $[A, B]$ to a function $\bar{f}(x)$ which is L.A.C. on $[A, B]$ and is such that $\partial\bar{f}/\partial x_k = D_{x_k}f$ almost everywhere.

We now choose an arbitrary cell $[a, b]$ interior to G. We shall show that there exists a set Z of $(n-1)$-dimensional measure zero in the cell $[a_k', b_k']$ such that $f_h(x_k', x_k)$ converges uniformly in x_k to $\bar{f}(x_k', x_k)$ for $a_k \leqq x_k \leqq b_k$, provided that x_k' is in $[a_k', b_k']$ and not in Z. It is clear that the lemma will follow from this as G may be covered by a denumerable number of such cells $[a, b]$.

Let us suppose first that $D_{x_k}f \geqq 0$ on a cell $[A, B]$ interior to G which contains $[a, b]$ in its interior, and let $\bar{f}(x)$ be L.A.C. on $[A, B]$ with $\partial\bar{f}/\partial x_k = D_{x_k}f$ almost everywhere on $[A, B]$ and $\bar{f}$ equivalent to f. Such a function $\bar{f}$ exists by virtue of Theorems 3.1 and 3.2. Then $\bar{f}(x_k', x_k)$ is monotone non-decreasing in x_k for any x_k' in $[A_k', B_k']$ which does not belong to a certain set Z_1 of $(n-1)$-dimensional measure zero. From the theorem of Fubini, it follows that, if c_k is not in a certain set of (linear) measure zero on $[A_k, B_k]$, then $f_h(x_k', c_k)$ tends to $\bar{f}(x_k', c_k)$ for all x_k' in $[A_k', B_k']$ which are not in a set $Z(c_k)$ of $(n-1)$-dimensional measure zero; this set may depend on c_k. We may then choose from these c_k a denumerable set S which contains $\bar{a}_k$ and $\bar{b}_k$ and is everywhere dense on the interval $\bar{a}_k \leqq x_k \leqq \bar{b}_k$, where $A_k < \bar{a}_k < a_k < b_k < \bar{b}_k < B_k$. We let Z consist of the set Z_1 and all the sets $Z(c_k)$ for c_k in S; Z is evidently of $(n-1)$-dimensional measure zero. Moreover, if x_k' is in $[a_k', b_k']$ and not in Z, then from Lemma 3.1 it follows that $f_h(x_k', x_k)$ and $\bar{f}(x_k', x_k)$ are all monotone non-decreasing in x_k for $\bar{a}_k \leqq x_k \leqq \bar{b}_k$ if h is sufficiently small, and from the above, we may conclude that $f_h(x_k', x_k)$ tends to $\bar{f}(x_k', x_k)$ for every x_k in S. From a well-known theorem on convergent sequences of continuous monotone functions, it follows that this convergence is uniform in x_k for each such x_k'.

If $f(x)$ does not have the property assumed in the above paragraph, we proceed as follows: Let $[A, B]$ be any cell interior to G which contains $[a, b]$ in its interior and let $\bar{f}$ be L.A.C. on the interior of $[A, B]$ with $\partial\bar{f}/\partial x_k = D_{x_k}f$ almost everywhere on $[A, B]$, and $\bar{f}$ equivalent to f on $[A, B]$. Let c_k, $A_k \leqq c_k \leqq B_k$, be chosen so that $\bar{f}(x_k', c_k)$ is integrable in x_k' over $[A_k', B_k']$. For each x_k' for which $\bar{f}(x_k', x_k)$ is A.C. in x_k, we define

$$f_1(x_k', \bar{x}_k) = \bar{f}(x_k', c_k) + \int_{c_k}^{\bar{x}_k} \frac{1}{2}\left[\left|\frac{\partial\bar{f}}{\partial x_k}\right| + \frac{\partial\bar{f}}{\partial x_k}\right] dx_k,$$

$$f_2(x_k', \bar{x}_k) = \int_{c_k}^{\bar{x}_k} \frac{1}{2}\left[\left|\frac{\partial\bar{f}}{\partial x_k}\right| - \frac{\partial\bar{f}}{\partial x_k}\right] dx_k;$$

we define $f_1 = \bar{f}, f_2 = 0$ elsewhere. Evidently $\bar{f} = f_1 - f_2$ on $[A, B]$, and it can be shown, as in the last part of the proof of Theorem 3.1, that f_1 and f_2 are L.A.C. on the interior of $[A, B]$ and are monotone non-decreasing in x_k for almost all

x'_k. As $f_h = \bar{f}_h$ and $\bar{f}_h = f_{1h} - f_{2h}$ for every h, the theorem follows in the present case from the preceding paragraph.

From the two preceding lemmas we have almost at once the two following theorems.

THEOREM 3.3. *A necessary and sufficient condition that a function $f(x)$ be essentially absolutely continuous in the variables $x_{k_1}, x_{k_2}, \cdots, x_{k_p}$ on G is that $f(x)$ be equal almost everywhere to a function f_0 linearly absolutely continuous in each of those variables on G.*

The sufficiency of the condition follows from the converse part of Theorem 3.1. On the other hand, the necessity is a consequence of Lemma 3.2; if $f(x)$ is E.A.C. in each of several variables, the function $f_0(x)$ of that lemma is evidently L.A.C. in each of those variables.

THEOREM 3.4. *A necessary and sufficient condition that a function $f(x)$ be E.A.C. on G in each of the variables $x_{k_1}, x_{k_2}, \cdots, x_{k_p}$ is that there exist functions $g_{k_1}, g_{k_2}, \cdots, g_{k_p}$ measurable on G and integrable on each closed cell interior to G, and for each such cell a sequence $\{f_q\}$, each member of class C′ on $[a, b]$, such that*

$$\lim_{q\to\infty} \int_a^b \left(|f - f_q| + \sum_{i=1}^{p} \left| g_{k_i} - \frac{\partial f_q}{\partial x_{k_i}} \right| \right) dx = 0.$$

When this condition is satisfied, we have $D_{x_{k_i}} f = g_{k_i}$ $(i = 1, \cdots, p)$ almost everywhere.

The necessity of the condition follows at once from Lemma 3.1 and Theorems 2.1 and 2.3.

To prove the sufficiency, we consider an arbitrary value k of k_i and an arbitrary cell $[a, b]$ in G. We then make use of a familiar theorem on mean convergence to choose a subsequence $\{f_{q_r}\}$ of $\{f_q\}$ so that

$$\lim_{r\to\infty} \int_{a'_k}^{b'_k} | f_{q_r}(x'_k, x_k) - f(x'_k, x_k) | \, dx'_k = 0$$

for almost all x_k on $a_k \leqq x_k \leqq b_k$. Thus we obtain

$$\lim_{r\to\infty} \int_{\alpha'_k}^{\beta'_k} f_{q_r}(x'_k, x_k) \, dx'_k = \int_{\alpha'_k}^{\beta'_k} f(x'_k, x_k) \, dx'_k$$

for $[\alpha'_k, \beta'_k]$ in $[a'_k, b'_k]$ and almost all x_k on $a_k \leqq x_k \leqq b_k$. But

$$\int_{\alpha'_k}^{\beta'_k} [f_{q_r}(x'_k, \beta_k) - f_{q_r}(x'_k, \alpha_k)] \, dx'_k = \int_\alpha^\beta \frac{\partial f_{q_r}}{\partial x_k} dx, \qquad a_k \leqq \alpha_k < \beta_k \leqq b_k,$$

and

$$\lim_{r\to\infty} \int_\alpha^\beta \frac{\partial f_{q_r}}{\partial x_k} dx = \int_\alpha^\beta g_k \, dx.$$

Thus

$$\int_{\alpha_k'}^{\beta_k'} [f(x_k', \beta_k) - f(x_k', \alpha_k)]\, dx_k' = \int_\alpha^\beta g_k\, dx$$

for $[\alpha_k', \beta_k']$ in $[a_k', b_k']$ and almost all β_k, α_k in $[a_k, b_k]$. From this it is evident that f is E.A.C. in x_k on $[a, b]$ and since k is an arbitrary value of k_i and $[a, b]$ arbitrary in G, it is easily shown that f is E.A.C. in x_{k_i} $(i = 1, 2, \cdots, p)$ on G.

The concluding assertion of the theorem is obvious.

For use later, we now note a possible modification of Theorem 3.4.

THEOREM 3.5. *Theorem* 3.4 *remains valid if the condition that the terms of the sequence* $\{f_q\}$ *be of class* C′ *on* $[a, b]$ *is replaced by the following:* f_q *is L.A.C. in* $x_{k_1}, x_{k_2}, \cdots, x_{k_p}$ *on the interior of* $[a, b]$ *and integrable, together with its partial derivatives with respect to those variables, on that cell.*

The necessity of the condition thus modified follows at once from the necessity of the original condition, and the sufficiency can be proved exactly as before.

4. **Functions of class $\mathfrak{P}$.** We pass now to the consideration of functions which are E.A.C. in all of the coördinate variables.

DEFINITION 4.1. A function $f(x)$ is *of class* $\mathfrak{P}$ on G if it is E.A.C. in each of the variables $x_1, x_2, \cdots, x_n$ on G.

DEFINITION 4.2. A function $f(x)$ is *of class* $\mathfrak{P}'$ on G if it is L.A.C. in each of the variables $x_1, x_2, \cdots, x_n$ on G.

As our introductory remarks indicated, our notion of a function of class $\mathfrak{P}$ is the same as Evans' concept of a potential function of its generalized derivatives. To be entirely precise, we may say that, for the case of a real-valued function of two variables, a function $f(x)$ is of class $\mathfrak{P}$ on G if and only if it is a potential function of its generalized derivatives on G.

In view of Theorem 3.3, we can state at once the following generalization of a theorem of Evans ([2], pp. 277–279).

THEOREM 4.1. *A function* $f(x)$ *is of class* $\mathfrak{P}$ *on* G *if and only if it is equal almost everywhere on* G *to a function of class* $\mathfrak{P}'$.

Similarly, Theorem 3.4 leads to a necessary and sufficient condition that a function be of class $\mathfrak{P}$. Actually, however, the following modification of that condition turns out to be more useful.

THEOREM 4.2. *A necessary and sufficient condition that a function* $f(x)$ *be of class* $\mathfrak{P}$ *on* G *is that there exist functions* $g_1, g_2, \cdots, g_n$ *measurable on* G *and absolutely integrable on every closed cell* $[a, b]$ *interior to* G, *and for each such cell a sequence* $\{f_q\}$, *each member satisfying a uniform Lipschitz condition on* $[a, b]$, *such that*

$$\lim_{q\to\infty} \int_a^b \left(|f - f_q| + \sum_{k=1}^n \left| g_k - \frac{\partial f_q}{\partial x_k} \right| \right) dx = 0.$$

The necessity of the condition follows at once from Lemma 3.1 and Theorems 2.1 and 2.3. The sufficiency of the condition, on the other hand, can be deduced immediately from Theorem 3.5, when we note that the Lipschitz condition on f_q ($q = 1, 2, \cdots$) implies that it satisfies the condition of that theorem.

The following theorem is almost obvious but because of its fundamental importance, deserves formal statement.

THEOREM 4.3. *Let $f_1(x)$ and $f_2(x)$ be of class $\mathfrak{P}$ on G and suppose that $D_{x_i}f_1 = D_{x_i}f_2$ almost everywhere on G ($i = 1, \cdots, n$). Then*

$$f_1(x) = f_2(x) + f^*(x) + C$$

on G, where $f^*(x) = 0$ almost everywhere on G and C is a constant. In other words two functions of class $\mathfrak{P}$ with essentially the same generalized derivatives differ at most by a constant plus a null function.

This theorem is a consequence of Theorem 2.2 and Lemma 3.1. For, from Theorem 2.2, we have

$$\lim_{h\to 0} f_{1h}(x) = f_1(x), \qquad \lim_{h\to 0} f_{2h}(x) = f_2(x)$$

almost everywhere on G. From Lemma 3.1, it follows that f_{1h} and f_{2h} are of class C′ on G_h and that their partial derivatives all coincide on G_h. Thus, we have

$$f_{1h}(x) = f_{2h}(x) + c_h$$

on G_h. Since f_1 and f_2 are integrable on each cell in G and since, by Theorem 2.3,

$$\lim_{h\to 0} \int_a^b |f_{1h}(x) - f_1(x)|\,dx = \lim_{h\to 0} \int_a^b |f_{2h}(x) - f_2(x)|\,dx = 0$$

for each cell interior to G, it follows that c_h tends to a limit c as $h \to 0$ and that

$$\int_a^b [f_1(x) - f_2(x)]\,dx = c\cdot\text{meas}\,([a, b])$$

for each cell in G. This establishes the theorem.

Some time after the introduction by Evans of the notion of a potential function of its generalized derivatives, Tonelli introduced the concept of an absolutely continuous function of two variables [6, 7]; and subsequently Evans discussed in some detail the relationship between his work and that of Tonelli [4]. It is worth while for us here to set down in our own terminology the salient points of that relationship.

We have first to introduce the natural generalization to the case of n variables of Tonelli's definition for the case of two variables.

DEFINITION 4.3. Let $f(x)$ be a real-valued function defined on a closed cell $[a, b]$. Then $f(x)$ is *absolutely continuous in the sense of Tonelli* (*A.C.T.*) on $[a, b]$ if it is continuous there, is absolutely continuous in x_k on $a_k \leq x_k \leq b_k$ for almost all fixed values of x_k' on $[a_k', b_k']$ ($k = 1, 2, \cdots, n$) and if it has the following

property: the variation $V_{a_k}^{b_k}[f(x_k', x_k)]$ of $f(x)$ as a function of x_k on the interval $a_k \leqq x_k \leqq b_k$, is integrable with respect to x_k' over $[a_k', b_k']$; that is to say, the integral

$$\int_{a_k'}^{b_k'} V_{a_k}^{b_k}[f(x_k', x_k)]\, dx_k'$$

exists ($k = 1, 2, \cdots, n$).

A real-valued function defined on G is said to be absolutely continuous in the sense of Tonelli (A.C.T.) on G if it is A.C.T. on every closed cell interior to G. A complex-valued function is said to be A.C.T. on G if its real and imaginary parts have that property.

To facilitate the comparison which we wish to make, we introduce also

DEFINITION 4.4. A function $f(x)$ defined on G is of class $\mathfrak{P}''$ on G if it is a continuous function of class $\mathfrak{P}$ on G.

In view of Theorem 4.1, it is evident that we have $\mathfrak{P}'' \subset \mathfrak{P}' \subset \mathfrak{P}$. Moreover, it is easily shown that each of these relations subsists in the strict sense, provided we exclude the case $n = 1$; in that case, the only reasonable interpretation of Definition 3.3 leads to the relation $\mathfrak{P}'' = \mathfrak{P}'$.

The relationship between the ideas developed in the preceding section and the concept of absolute continuity in the sense of Tonelli can now be adequately indicated in a formal statement.

THEOREM 4.4. *A function $f(x)$ is A.C.T. on G if and only if it is of class $\mathfrak{P}''$ there. Thus every continuous function of class $\mathfrak{P}$ on G is* A.C.T., *and conversely.*

For the proof, we refer the reader to a previously cited memoir of Evans [4]; although the argument there deals only with functions of two variables, the demonstration for the general case can be carried through in entirely analogous fashion.

Before we proceed, it is convenient to state for later use the following obvious theorem, omitting proof.

THEOREM 4.5. *Let $f(x, y)$ be of class $\mathfrak{P}$ ($\mathfrak{P}'$ or $\mathfrak{P}''$) in the $n + m$ variables $x_1, x_2, \cdots, x_n, y_1, y_2, \cdots, y_m$ on an open set G in $(n + m)$-dimensional space. Let $G(x)$ denote the set of values of y for which the point (x, y) is in G. Then, for almost all values of x for which $G(x)$ is not empty $f(x, y)$ is of class $\mathfrak{P}$ ($\mathfrak{P}'$ or $\mathfrak{P}''$) on $G(x)$.*

To conclude this section, we observe that each of the classes $\mathfrak{P}$, $\mathfrak{P}'$, and $\mathfrak{P}''$, in both the real and complex cases, is a module. That is to say, in both cases, each class is closed with respect to ordinary addition; in the case of real-valued functions, each is closed with respect to scalar multiplication by real numbers, and in the complex case, each is closed with respect to multiplication by complex numbers. In addition, we may note that in the complex case, the complex conjugate of a function of class $\mathfrak{P}$ ($\mathfrak{P}'$ or $\mathfrak{P}''$) belongs to the same class.

Finally, we point out that the class obtained by reducing $\mathfrak{P}$ modulo the class of functions equal almost everywhere to zero on G is also a module, and that

each residue class so obtained contains a function of class $\mathfrak{P}'$. We emphasize, however, that the latter function is not unique; if $n \geqq 2$, and f is of class $\mathfrak{P}'$, any function obtained from f by changing its value at a finite number of points is also of class $\mathfrak{P}'$.

5. **Functions of class $\mathfrak{P}_\alpha$.** In Theorems 3.4, 3.5, and 4.2, we have discussed the matter of approximating to functions E.A.C. in each of several variables and to functions of class $\mathfrak{P}$, making use of mean convergence of order 1 on closed cells interior to G. We turn now to similar and related considerations, dealing, however, with mean convergence of arbitrary order $\alpha \geqq 1$.

To begin, we introduce

DEFINITION 5.1. Let $f(x)$ be of class $\mathfrak{P}$ on G. Let R be an arbitrary measurable set interior to G. We define $D_\alpha(f, R)$, $\alpha \geqq 1$, by the equation

$$D_\alpha(f, R) = \int_R \left(\sum_{k=1}^n |D_{x_k} f|^2 \right)^{\frac{1}{2}\alpha} dx$$

if the integral involved is finite; otherwise, we write $D_\alpha(f, R) = \infty$. Similarly, we define $\bar{D}_\alpha(f, R)$, $\alpha \geqq 1$, by the equation

$$\bar{D}_\alpha(f, R) = D_\alpha(f, R) + \int_R |f|^\alpha dx,$$

if the integrals involved are finite; otherwise, we write $\bar{D}_\alpha(f, R) = \infty$.

DEFINITION 5.2. Let f be of class $\mathfrak{P}$ ($\mathfrak{P}'$ or $\mathfrak{P}''$) on G. If $\bar{D}_\alpha(f, G)$ is finite, f is said to be of class $\mathfrak{P}_\alpha$ ($\mathfrak{P}'_\alpha$ or $\mathfrak{P}''_\alpha$) on G.

We note in passing that Theorem 4.5 remains true, if we substitute for $\mathfrak{P}$ ($\mathfrak{P}'$ or $\mathfrak{P}''$) the class $\mathfrak{P}_\alpha$ ($\mathfrak{P}'_\alpha$ or $\mathfrak{P}''_\alpha$).

In addition, it is readily verified that the remarks of the two concluding paragraphs of §4 apply equally well to functions of class $\mathfrak{P}_\alpha$ ($\mathfrak{P}'_\alpha$ or $\mathfrak{P}''_\alpha$) as to functions of class $\mathfrak{P}$ ($\mathfrak{P}'$ or $\mathfrak{P}''$). Thus the class obtained by identifying equivalent functions in the class $\mathfrak{P}_\alpha$ is a linear metric space, since it is readily verified that in this class the function $\|f\| = [\bar{D}_\alpha(f, G)]^{1/\alpha}$ has the requisite properties of a norm. Later, we shall see that this space is even complete and is therefore a Banach space, which is a Hilbert space in the case $\alpha = 2$.

We proceed now to an elementary discussion of the relation between the quantities $D_\alpha(f, R)$ and $\bar{D}_\alpha(f, R)$.

THEOREM 5.1. *If $f(x)$ is of class $\mathfrak{P}$ ($\mathfrak{P}'$ or $\mathfrak{P}''$) on the open cell R with edges parallel to the coördinate axes and $D_\alpha(f, R)$ is finite, then $\bar{D}_\alpha(f, R)$ is finite and thus f is of class $\mathfrak{P}_\alpha$ ($\mathfrak{P}'_\alpha$ or $\mathfrak{P}''_\alpha$) on R.*

We observe first that the interpolatory statement concerning functions of class $\mathfrak{P}''$ is true if that concerning functions of class $\mathfrak{P}'$ is true, since a function of class $\mathfrak{P}'$ ($\mathfrak{P}'_\alpha$) is of class $\mathfrak{P}''$ ($\mathfrak{P}''_\alpha$) if and only if it is continuous. Thus, since every function of class $\mathfrak{P}$ is equivalent to a function of class $\mathfrak{P}'$, we can prove the

theorem in its entirety, by showing merely that $\int_R |f|^\alpha\, dx$ exists when f is of class $\mathfrak{P}'$ on R and $D_\alpha(f, R)$ is finite.

The assertion in question is obviously true for $n = 1$, and we prove it for the general case by induction. Taking $n = k$, we assume then that the theorem is true for $n = k - 1$. Hence, if $D_\alpha(f, R)$ is finite, we have $\int_{a_k'}^{b_k'} |f(x_k', c_k)|^\alpha\, dx_k'$ finite for almost all values of c_k on the interval $a_k \leqq x_k \leqq b_k$. But, for almost all values of x_k' in $[a_k', b_k']$ we have

$$f(x_k', x_k) = f(x_k', c_k) + \int_{c_k}^{x_k} \frac{\partial f(x_k', x_k)}{\partial x_k}\, dx_k, \qquad a_k \leqq c_k \leqq b_k.$$

Thus, we have

$$|f(x_k', x_k)|^\alpha \leqq 2^{\alpha-1}\left[|f(x_k', c_k)|^\alpha + \left|\int_{c_k}^{x_k} \frac{\partial f}{\partial x_k}\, dx_k\right|^\alpha\right],$$

since $(|a| + |b|)^\alpha \leqq 2^{\alpha-1}(|a|^\alpha + |b|^\alpha)$, and hence

$$|f(x_k', x_k)|^\alpha \leqq 2^{\alpha-1}|f(x_k', c_k)|^\alpha + 2^{\alpha-1}\cdot(b_k - a_k)^{\alpha-1}\int_{a_k}^{b_k}\left|\frac{\partial f}{\partial x_k}\right|^\alpha dx_k$$

for almost all x_k' in $[a_k', b_k']$. Hence, since the right member of the last inequality is clearly integrable over $[a, b]$ for suitably chosen c_k, it follows that $\int_a^b |f|^\alpha\, dx$ exists, and the proof is complete.

In the following paper, C. B. Morrey will establish Theorem 5.1 for a more general open set than a cell. For the present we confine ourselves to the preceding theorem and the one following.

THEOREM 5.2. *Let $f(x)$ be of class $\mathfrak{P}$ ($\mathfrak{P}'$ or $\mathfrak{P}''$) on the open set G and let $D_\alpha(f, G)$ be finite. Then $\bar{D}_\alpha(f, R)$ is finite for every bounded measurable set R whose closure belongs to G.*

The proof is immediate, since we can cover the closure of R by a finite number of cells $[a, b]$ each interior to G and then apply Theorem 5.1.

LEMMA 5.1. *Let $f(x)$ be of class $\mathfrak{P}$ on G and of class $\mathfrak{P}_\alpha$ on each bounded open set R whose closure belongs to G. Let $f_h(x)$ be the h-average of $f(x)$. Then*

$$\lim_{h\to 0} \bar{D}_\alpha(f - f_h, R) \equiv \lim_{h\to 0}\int_R\left[|f - f_h|^\alpha + \left(\sum_{k=1}^n \left|D_{x_k} f - \frac{\partial f_h}{\partial x_k}\right|^2\right)^{\frac{1}{2}\alpha}\right] dx = 0$$

for each such set R.

The lemma follows at once from Theorem 2.3 and Lemma 3.1.

THEOREM 5.3. *For $f(x)$ to be of class $\mathfrak{P}_\alpha$ on G, it is necessary and sufficient that the following conditions be satisfied:*

(i) *$\int_G |f|^\alpha dx$ is finite;*

(ii) *there exist functions* $g_1, \cdots, g_n$ *with* $\int_G |g_k|^\alpha dx$ *finite* $(k = 1, 2, \cdots, n)$ *and for each closed cell* $[a, b]$ *interior to* G *a sequence* $\{f_q\}$, *each member satisfying a uniform Lipschitz condition on that cell, such that*

$$\lim_{q\to\infty} \int_a^b \left(|f - f_q|^\alpha + \sum_{k=1}^n \left| g_k - \frac{\partial f_q}{\partial x_r}\right|^\alpha\right) dx = 0.$$

The necessity of (i) is obvious, as is the necessity of the condition that

$$\int_G |g_k|^\alpha dx \qquad (k = 1, 2, \cdots, n)$$

be finite, while the necessity of the remaining provisions of (ii) follows at once from Lemma 5.1.

With regard to the sufficiency, we note first that the equation with which the theorem concludes implies that

$$\lim_{r\to\infty} \int_{\alpha_k'}^{\beta_k'} f_{q_r} dx_k' = \int_{\alpha_k'}^{\beta_k'} f dx_k'$$

for all $[\alpha_k', \beta_k']$ in $[a_k', b_k']$ and almost all x_k on $a_k \leqq x_k \leqq b_k$ $(k = 1, 2, \cdots, n)$ and that

$$\lim_{r\to\infty} \int_\alpha^\beta \frac{\partial f_{q_r}}{\partial x_r} dx = \int_\alpha^\beta g_k dx \qquad (k = 1, 2, \cdots, n)$$

for all $[\alpha, \beta]$ in $[a, b]$, for a suitably chosen subsequence $\{f_{q_r}\}$ of $\{f_q\}$. Thus an argument exactly like that used in the proof of Theorem 3.4 implies that f is of class $\mathfrak{P}$ on G and that $D_{x_k} f = g_k$ almost everywhere on G. Thus, from the integrability conditions on f and g_k $(k = 1, 2, \cdots, n)$ it follows that f is of class $\mathfrak{P}_\alpha$ on G.

Paralleling Theorem 3.5, we note the following possible modification of Theorem 5.3:

THEOREM 5.4. *Theorem* 5.3 *remains valid if the condition that each term of the sequence* $\{f_q\}$ *satisfy a uniform Lipschitz condition on* $[a, b]$ *is replaced by the condition that* f_q *be of class* $\mathfrak{P}_\alpha'$ *on the interior of* $[a, b]$.

The proof requires scarcely any modification, as the reader will readily perceive.

It is now of some importance to observe that when f is of class $\mathfrak{P}''$, we can approximate to f interior to G in a much stricter sense than is indicated by Theorem 5.3.

THEOREM 5.5. *For* $f(x)$ *to be of class* $\mathfrak{P}_\alpha''$ *on* G, *it is necessary and sufficient that the condition* (i) *of Theorem* 5.3 *be satisfied and that there exist functions* $g_1, g_2, \cdots, g_n$ *with* $\int_G |g_k|^\alpha dx$ $(k = 1, 2, \cdots, n)$ *finite, and for each closed cell* $[a, b]$ *interior*

to G a sequence $\{f_q\}$, each member satisfying a uniform Lipschitz condition on that cell, converging uniformly to f there, with

$$\lim_{q\to\infty} \int_a^b \sum_{k=1}^n \left| g_k - \frac{\partial f_q}{\partial x_k} \right|^\alpha dx = 0.$$

To establish the necessity of the condition, we need only supplement the proof of Theorem 5.3 with the observation that when f is continuous on G, $f_h = (2h)^{-n} \int_{x-h}^{x+h} f\,dx$ converges to it at every point of G, the convergence being uniform on cells interior to G. The sufficiency of the condition is quite evident.

THEOREM 5.6. *Let $\{f_q\}$ be a sequence of functions of class $\mathfrak{P}_\alpha$ on G such that*

$$\lim_{p,q\to\infty} \bar{D}_\alpha(f_p - f_q, G) = 0.$$

Then there exists a function f of class $\mathfrak{P}'_\alpha$ on G such that

$$\lim_{q\to\infty} \bar{D}_\alpha(f - f_q, G) = 0.$$

We note first that Theorem 4.1 permits us to assume that the functions f_q $(q = 1, 2, \cdots)$ are of class $\mathfrak{P}'_\alpha$, since every function of class $\mathfrak{P}_\alpha$ is equivalent to such a function. Next we observe that the hypothesis of the theorem implies that the sequences $\{f_q\}$, $\{\partial f_q/\partial x_k\}$ $(k = 1, 2, \cdots, n)$ converge in the mean of order α on G. Hence the respective limits $f, g_1, g_2, \cdots, g_n$, in that sense, of these sequences exist. Furthermore, if $f, g_1, g_2, \cdots, g_n$, $\{f_q\}$ have the same meanings as in Theorem 5.3, it is clear that the conditions of that theorem, with the modification noted in Theorem 5.4, are satisfied. Thus f is of class $\mathfrak{P}_\alpha$ on G, $D_{x_k} f = g_k$ almost everywhere $(k = 1, 2, \cdots, n)$ and

$$\lim_{q\to\infty} \bar{D}_\alpha(f - f_q, G) = 0.$$

Consequently, with the observation that f is equivalent to a function of class $\mathfrak{P}'_\alpha$, the proof is complete.

Thus, recalling the remarks following Definition 5.2, we perceive that $\mathfrak{P}_\alpha$, with equivalent functions identified, constitutes a Banach space which is a Hilbert space if $\alpha = 2$.

COROLLARY. *Let $\{f_q\}$ be a sequence of functions of class $\mathfrak{P}_\alpha$ on G such that*

$$\lim_{p,q\to\infty} \bar{D}_\alpha(f_p - f_q, [a, b]) = 0$$

for every closed cell $[a, b]$ interior to G. Then there exists a function f of class $\mathfrak{P}'$ on G and of class $\mathfrak{P}'_\alpha$ on every open set R with closure in G such that

$$\lim_{q\to\infty} \bar{D}_\alpha(f - f_q, R) = 0.$$

It is natural now to consider possible restrictions on the sequence $\{f_q\}$ of the preceding corollary, or on the sequence $\{f_q\}$ of Theorem 5.6, which would serve to insure the existence of a continuous limit f.

We conclude by considering one condition of this sort which is especially useful.

THEOREM 5.7. *Let the terms of the sequence* $\{f_q\}$ *of the preceding corollary be of class* $\mathfrak{P}''$ *and equicontinuous on each cell* $[a, b]$ *in* G. *Then* $\{f_q\}$ *converges uniformly on each such cell to a function of class* $\mathfrak{P}''$.

We consider an arbitrary closed cell $[a, b]$ in G and choose a subsequence $\{f_{q_p}\}$ of $\{f_q\}$, converging almost everywhere on $[a, b]$. Invoking the equicontinuity, we can then conclude first that the subsequence converges uniformly on $[a, b]$, and secondly that the entire sequence does. Since the limit in the sense of this convergence is clearly equal almost everywhere on G to the limit described in the corollary, it follows that it is of class $\mathfrak{P}''$ on G.

BIBLIOGRAPHY

1. H. E. BRAY, *Proof of a formula for an area*, Bulletin of the American Mathematical Society, vol. 29(1923), pp. 264–270.
2. G. C. EVANS, *Fundamental Points of Potential Theory*, Rice Institute Pamphlets, vol. 7(1920), no. 4, pp. 252–329, especially pp. 274–285.
3. G. C. EVANS, *Note on a theorem of Bôcher*, American Journal of Mathematics, vol. 50(1928), pp. 123–126.
4. G. C. EVANS, *Complements of potential theory*, Part II, American Journal of Mathematics, vol. 55(1933), pp. 29–49, especially pp. 42–46.
5. C. B. MORREY, JR., *A class of representations of manifolds*, Part I, American Journal of Mathematics, vol. 55(1933), pp. 683–707, especially p. 687.
6. L. TONELLI, *Sulla quadratura delle superficie*, Atti della Reale Accademia dei Lincei, ser. 6, vol. 3(1926), pp. 633–638.
7. L. TONELLI, Comptes Rendus Hebd. des Séances de l'Académie des Sciences, vol. 182(1926), pp. 1198–1200.
8. C. J. DE LA VALLÉE POUSSIN, *Intégrales de Lebesgue, Fonctions d'Ensembles, Classes de Baire*, Paris, 1916, pp. 59–61.

ARMOUR INSTITUTE OF TECHNOLOGY.

FUNCTIONS OF SEVERAL VARIABLES AND ABSOLUTE CONTINUITY, II

By C. B. Morrey, Jr.

In Part I of this paper, by J. W. Calkin [2],[1] functions of class $\mathfrak{P}$, $\mathfrak{P}'$, $\mathfrak{P}''$, $\mathfrak{P}_\alpha$, $\mathfrak{P}'_\alpha$, and $\mathfrak{P}''_\alpha$ were defined and various theorems were proved concerning these functions. The object of this part is to carry forward the study of these functions to the point where they may be used in the study of standard problems of analysis, particularly those problems in which the authors are interested. Part I of this paper comprises §§1–5 and the present part contains §§6–9 inclusive. We shall not refer to Part I by name but will merely refer to certain theorems, lemmas, and definitions therein, such as Theorem 3.5, etc.; in other words we shall regard the two parts as a single unit.

Certain notations were introduced in §1 which simplify the discussion considerably. We shall continue the use of these notations and shall present them before proceeding with the discussion. The n-tuples $(x_1, \cdots, x_n)$, $(\xi_1, \cdots, \xi_n)$, etc., are denoted by the single letters x, ξ, etc. The closed cell $a_i \leqq x_i \leqq b_i$ $(i = 1, \cdots, n)$ is denoted by $[a, b]$ and the open cell by (a, b). The functional notation $f(x_1, \cdots, x_n)$ is abbreviated to $f(x)$ and the Lebesgue integrals of $f(x)$ over (a, b) and a set E are denoted respectively by

$$\int_a^b f(x)\,dx, \qquad \int_E f(x)\,dx.$$

It is frequently desirable to consider the behavior of $f(x)$ with reference to a particular variable x_k or the $n - 1$ variables $(x_1, \cdots, x_{k-1}, x_{k+1}, \cdots, x_n)$. In such a case, we write x'_k for $(x_1, \cdots, x_{k-1}, x_{k+1}, \cdots, x_n)$, (x'_k, x_k) for x, and $f(x'_k, x_k)$ for $f(x)$. Thus, if the coördinates x'_k are fixed with $x'_k = a'_k$, $f(a'_k, x_k)$ denotes the function $f(a_1, \cdots, a_{k-1}, x_k, a_{k+1}, \cdots, a_n)$ of the single variable x_k. Similarly, if x_k is fixed with $x_k = a_k$, $f(x'_k, a_k)$ denotes the function $f(x_1, \cdots, x_{k-1}, a_k, x_{k+1}, \cdots, x_n)$ of the $n - 1$ variables $(x_1, \cdots, x_{k-1}, x_{k+1}, \cdots, x_n) = x'_k$. We denote the projection of the closed cell $[a, b]$ on $x_k = 0$ by $[a'_k, b'_k]$ and that of (a, b) by (a'_k, b'_k). The integral

$$\int_{a'_k}^{b'_k} f(x'_k, a_k)\,dx'_k$$

denotes the $(n - 1)$-dimensional integral of $f(x'_k, a_k)$ over the cell (a'_k, b'_k), i.e., over the $(n - 1)$-cell $a_1 < x_1 < b_1, \cdots, a_{k-1} < x_{k-1} < b_{k-1}, a_{k+1} < x_{k+1} <$

Received September 12, 1939.

[1] Numbers in brackets refer to the bibliography.

$b_{k+1}, \cdots, a_n < x_n < b_n$. We may also wish to write the n-dimensional integrals as iterated integrals

$$\int_a^b f(x)\,dx = \int_{a_k}^{b_k}\int_{a_k'}^{b_k'} f(x_k', x_k)\,dx_k\,dx_k' = \int_{a_k'}^{b_k'}\int_{a_k}^{b_k} f(x_k', x_k)\,dx_k'\,dx_k$$

in either order as the occasion warrants. We usually consider our functions to be real but they may be taken to be complex valued if certain trivial modifications are made.

We now recall the fundamental definitions given in Part I. A function $z(x)$ is said to be *essentially absolutely continuous* (*E.A.C.*) *in* x_k on a region G if (1) $z(x)$ is summable on each cell of G and (2) there exists a function $v_k(x)$ satisfying (1) such that the formula

$$\text{(A)}\qquad \int_{a_k'}^{b_k'} [z(x_k', b_k) - z(x_k', a_k)]\,dx_k' = \int_a^b v_k(x)\,dx$$

holds for almost every cell $[a, b]$ in G (i.e., for each cell $[a, b]$ in G for which the point $(a_1, \cdots, a_n, b_1, \cdots, b_n)$ does not lie in a $2n$-dimensional set of measure zero). A function $z(x)$ is said to be *linearly absolutely continuous* (*L.A.C.*) *in* x_k on G if (1) $z(x)$ is summable on each cell of G, (2) $z(\bar{x}_k', x_k)$ is A.C. in x_k on each closed segment of the line $x_k' = \bar{x}_k'$ which is in G provided that $\bar{x}_k'$ is not in a certain set of measure zero, and (3) $\partial z/\partial x_k$ (which exists almost everywhere from (1) and (2)) satisfies (1). If $z(x)$ is E.A.C. in x_k, we define the *generalized derivative* ([4], pp. 274–275) $D_{x_k}z$ *of* z *with respect to* x_k as the Lebesgue derivative ([11], pp. 59–61) of the set function $\int_e v_k(x)\,dx$; this exists and is equal to $v_k(x)$ almost everywhere. A function $z(x)$ is *of class* $\mathfrak{P}$ on G if it is E.A.C. in x_k for each k $(1 \leq k \leq n)$; it is *of class* $\mathfrak{P}'$ on G if it is L.A.C. in x_k on G for each k; and it is *of class* $\mathfrak{P}''$ on G if it is of class $\mathfrak{P}$ and continuous on G. The function z is *of class* $\mathfrak{P}_\alpha$ *or* $\mathfrak{P}_\alpha''$ on G if it is of class $\mathfrak{P}$ or $\mathfrak{P}''$ (respectively) on G with $|z|^\alpha$ and $|D_{x_i}z|^\alpha$ summable on G $(i = 1, \cdots, n)$; and z is *of class* $\mathfrak{P}_\alpha'$ on G if it is of class $\mathfrak{P}'$ on G with $|z|^\alpha$ and $|\partial z/\partial x_i|^\alpha$ summable on G $(i = 1, \cdots, n)$.

In §§4 and 5, the classes of functions of classes $\mathfrak{P}$ and $\mathfrak{P}''$ were identified respectively with the class of "potential functions of their generalized derivatives" of G. C. Evans ([4], p. 274; [5]) and the class of functions absolutely continuous in the sense of Tonelli ([10]; [3], pp. 42–46). The following are the further essential facts proved in §§4 and 5:

(1) any function equivalent to a function of class $\mathfrak{P}$ or $\mathfrak{P}_\alpha$ is also of class $\mathfrak{P}$ or $\mathfrak{P}_\alpha$ and possesses the same generalized derivatives defined at the same points;

(2) any two functions of class $\mathfrak{P}$ which have the same generalized derivatives almost everywhere differ by a constant and a null function;

(3) any function of class $\mathfrak{P}'$ $(\mathfrak{P}_\alpha')$ is of class $\mathfrak{P}$ $(\mathfrak{P}_\alpha)$, its partial and generalized derivatives coincide almost everywhere, and formulas (A) hold for each cell in G if we take $v_k(x) = \partial z/\partial x_k$ $(k = 1, \cdots, n)$;

(4) any function of class $\mathfrak{P}$ is equivalent to a function of class $\mathfrak{P}'$;

(5) necessary and sufficient conditions that a function $f(x)$ be of class $\mathfrak{P}$ on G are that f be summable on each cell in G, that there exist functions $g_1, \cdots, g_n$ with this property and that, for each cell $[a, b]$ in G, there exists a sequence $\{f_n(x)\}$, each function of class C′ or satisfying a uniform Lipschitz condition on $[a, b]$ such that f_n and $\partial f_n/\partial x_i$ tend in the mean of order 1 to f and g_i respectively $(i = 1, \cdots, n)$; if f is of class $\mathfrak{P}''$, f_n may be chosen to converge uniformly to f and, if f is of class $\mathfrak{P}_\alpha$, the mean convergence may be taken to be of order α; and

(6) the space $\mathfrak{P}_\alpha$ ($\alpha \geqq 1$) of classes of equivalent functions of class $\mathfrak{P}_\alpha$ is a Banach space if we define

$$\|z\|^\alpha = \bar{D}_\alpha(z, G) = \int_G |z|^\alpha dx + D_\alpha(z, G), \qquad D_\alpha(z, G) = \int_G \left[\sum_{i=1}^n D_{x_i}^2 z\right]^{\frac{1}{2}\alpha} dx.$$

In §6, we present a proof of a theorem of G. C. Evans ([4], p. 282) concerning the invariance of the classes $\mathfrak{P}$ and $\mathfrak{P}''$ under certain transformations of coördinates. In §7, we define a rather general class of regions and discuss the behavior of functions of class $\mathfrak{P}_\alpha$ on the boundary of such regions. In §8, weak convergence in the space $\mathfrak{P}_\alpha$ is discussed and many important theorems are proved. Finally §9 gives a brief discussion of functions of class $\mathfrak{P}_\alpha$ defined on arbitrary bounded regions. The developments in §9 are closely related to those in volume 2, Chapter VII, §1 of Courant-Hilbert's *Methoden der mathematischen Physik* (Berlin, 1937).

By a *region*, we mean an open connected set. If G is a region, G^* denotes its boundary and $\bar{G}$ its closure. We shall frequently use the notation $C(P, r)$ or $C(x, r)$ to denote the sphere with center at P or x and radius r. The coördinates of points P or P_0 may be denoted by x_P or $x_0 = (x_{10}, \cdots, x_{n0})$ or $(x_{01}, \cdots, x_{0n})$; x_1, x_2, etc. shall stand for $(x_{1,1}, \cdots, x_{n,1})$, $(x_{1,2}, \cdots, x_{n,2})$, etc.

6. A theorem on change of variables. In this section, we wish to give a simple proof of a theorem of G. C. Evans ([4], p. 282) concerning changes of variable in functions of class $\mathfrak{P}$. He has proved this theorem for transformations of class C′ in two variables. We shall extend his results to the somewhat more general transformations of class K, defined below, in n dimensions.

We now give the following definitions.

DEFINITION 6.1. A transformation $x = x(y)$ of a set T in the y-space into a set S in the x-space is said to be *of class* C′ if it is 1-1 and continuous and if the functions $x_i(y)$ and $y_i(x)$ (of the inverse) are of class C′ on T and S respectively.[2]

[2] A function $\varphi(x)$ is of class C′ on a set S if there exist functions $a_i(x)$, continuous on S, such that

$$\varphi(\xi) = \varphi(x) + \sum_{i=1}^n a_i(x)(\xi_i - x_i) + r\cdot\epsilon(x, \xi),$$

$$\lim_{\xi\to x} \epsilon(x, \xi) = 0, \qquad \xi \in S, \qquad x \in S, \qquad r^2 = \sum_{i=1}^n (\xi_i - x_i)^2,$$

$\epsilon(x, \xi)$ being continuous in ξ on S for each x on S.

The transformation is also *regular* if the functions $x_i(y)$ and $y_i(x)$ all satisfy uniform Lipschitz conditions on T and S respectively.

DEFINITION 6.2. The above transformation $x = x(y)$ of T into S is said to be *of class* K if it is 1-1 and continuous and if the functions $x_i(y)$ and $y_i(x)$ satisfy uniform Lipschitz conditions on each bounded closed subset of T and S respectively. The transformation is also *regular* if the functions $x_i(y)$ and $y_i(x)$ satisfy uniform Lipschitz conditions on the whole T and S respectively.

Clearly a transformation of class C′ is also of class K and a regular transformation of class C′ is a regular transformation of class K. It is also clear that the totality of transformations in any one of these four classes forms a group.

Rademacher [8] has considered transformations of class K in two dimensions at great length and many other authors (including W. H. Young, H. E. Bray, T. Radó, E. J. McShane, and C. B. Morrey) have studied more general transformations. Rademacher has outlined a method of carrying over his results to n dimensions. Accordingly, we shall simply quote Rademacher's results for n dimensions. The first result [8] is the following:

LEMMA 6.1. *If $f(x)$ satisfies a uniform Lipschitz condition on the open set G, it possesses a total differential almost everywhere on G; i.e., if x_0 is not in a certain set of measure zero in G, the partial derivatives $(\partial f/\partial x_i)_{x_0}$ all exist and*

$$f(x) = f(x_0) + \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}\right)_{x_0} (x_i - x_{i,0}) + r\cdot\epsilon(x), \qquad r^2 = \sum_{i=1}^{n} (x_i - x_{i,0})^2,$$

where $\epsilon(x)$ is continuous in G and vanishes at $x = x_0$.

Before stating further results, we wish to introduce a simplified notation for certain Jacobians which is in line with our previous notations. Let $x = x(y)$ be a transformation of a set in the y-space into a set in the x-space. If all the partial derivatives $\partial x_i/\partial y_j$ exist at some point we define

$$\text{(6.1)} \qquad \frac{dx}{dy} = \frac{\partial(x_1, \cdots, x_n)}{\partial(y_1, \cdots, y_n)}, \qquad \frac{\partial x_i'}{\partial y_j'} = \frac{\partial(x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n)}{\partial(y_1, \cdots, y_{j-1}, y_{j+1}, \cdots, y_n)},$$

where the symbols on the right have their classical significance. Clearly we have

$$\text{(6.2)} \qquad \sum_{i=1}^{n} (-1)^{i+j} \frac{\partial x_i}{\partial y_j} \frac{\partial x_i'}{\partial y_k'} = \sum_{i=1}^{n} (-1)^{i+j} \frac{\partial x_j}{\partial y_i} \frac{\partial x_k'}{\partial y_i'} = \delta_{j,k} \frac{dx}{dy},$$

where $\delta_{j,k}$ has its usual significance as the Kronecker delta.

We now sum up the known [8] or easily deducible results concerning regular transformations of class K in the following lemma:

LEMMA 6.2. *Let $x = x(y)$ be a regular transformation of class* K *of an open set T in the y-space into an open set S in the x-space. Then*

(1) *measurable subsets of S and T correspond;*

(2) *if $s \subset S$ and $t \subset T$ are corresponding measurable sets, then*

$$m(s) = \int_t \left|\frac{dx}{dy}\right| dy, \qquad m(t) = \int_s \left|\frac{dy}{dx}\right| dx;$$

(3) *dx/dy has the same sign at any two points y where all the $x_i(y)$ have total differentials (i.e., $dx/dy > 0$ almost everywhere or $dx/dy < 0$ almost everywhere);*

(4) *if $f(x)$ is summable on S and if $g(y) = f[x(y)]$, then $g(y)$ and $g(y)\cdot|dx/dy|$ are both summable on T and if $s \subset S$ and $t \subset T$ are corresponding measurable sets, then*

$$\int_s f(x)\,dx = \int_t g(y)\cdot\left|\frac{dx}{dy}\right|dy, \qquad \int_t |g(y)|\,dy \leqq K\cdot\int_s |f(x)|\,dx,$$

where K depends only on the transformation; and

(5) *if x_0 and y_0 correspond, then regular families (in the sense of Lebesgue;* see [11], pp. 59–61) *of sets about x_0 and y_0 correspond.*

If all of the functions $x_i(y)$ possess total differentials at y_0 and if $x_0 = x(y_0)$, then

(6) *the same is true of the inverse functions $y_i(x)$ at x_0 and*

$$\left(\frac{dy}{dx}\right)_{x_0}\cdot\left(\frac{dx}{dy}\right)_{y_0} = 1, \qquad \left(\frac{dx}{dy}\right)_{y_0}\cdot\left(\frac{\partial y_i}{\partial x_j}\right)_{x_0} = (-1)^{i+j}\left(\frac{\partial x_j'}{\partial y_i'}\right)_{y_0};$$

(7) *if $f(x)$ possesses a total differential at x_0, then $g(y)$ possesses a total differential at $y = y_0$ and*

$$\left(\frac{\partial g}{\partial y_k}\right)_{y_0} = \sum_{i=1}^{n}\left(\frac{\partial f}{\partial x_i}\right)_{x_0}\cdot\left(\frac{\partial x_i}{\partial y_k}\right)_{y_0} \qquad (k = 1, \cdots, n);$$

and

(8) *if $f(y)$ is summable on T and the Lebesgue derivative* ([11], pp. 59–61) *of the set function $\int_t f(y)\,dy$ exists at y_0 and is equal to $f(y_0)$, and if $\{t\}$ is a regular family of sets at y_0 and s corresponds to t ($\{s\}$ then forming a regular family by* (5)), *we have*

$$\lim_{t\to y_0}\frac{m(s)}{m(t)} = \left|\frac{dx}{dy}\right|_{y_0},$$

$$\lim_{t\to y_0}\frac{1}{m(t)}\int_t f(y)\cdot R\left[\frac{\partial x_i}{\partial y_j}\right]\cdot\left|\frac{dx}{dy}\right|dy = f(y_0)\cdot R\left[\left(\frac{\partial x_i}{\partial y_j}\right)_{y_0}\right]\cdot\left|\frac{dx}{dy}\right|_{y_0},$$

where $R[\partial x_i/\partial y_j]$ is any rational function in the indicated partial derivatives which is defined at $y = y_0$.

We can now prove Evans' theorem on change of variable for transformations of class K.

THEOREM 6.1. *Let $f(x)$ be of class $\mathfrak{P}$ ($\mathfrak{P}''$) on the open set S and let $x = x(y)$ be a transformation of class K of the open set T into S. Then the function $g(y) = f[x(y)]$ is of class $\mathfrak{P}$ ($\mathfrak{P}''$) on T. Moreover, if all of the $x_i(y)$ possess total differentials at y_0, if $x_0 = x(y_0)$, and if all the generalized derivatives of f exist at x_0, then all the generalized derivatives of g exist at y_0 and are given by the usual formulas*

$$(D_{y_k}g)_{y_0} = \sum_{i=1}^{n}(D_{x_i}f)_{x_0}\cdot\left(\frac{\partial x_i}{\partial y_k}\right)_{y_0} \qquad (k = 1, \cdots, n). \tag{6.3}$$

If $f(x)$ is of class $\mathfrak{P}_\alpha$ ($\mathfrak{P}''_\alpha$) and the transformation is regular, then $g(y)$ is of class $\mathfrak{P}_\alpha$ ($\mathfrak{P}''_\alpha$) on T.

First, let $[c, d]$ be a cell of T which is small enough so that the closed set Σ corresponding to it lies in a cell $[a, b]$ of G; evidently our transformation is regular from $[c, d]$ to Σ. From Theorem 4.2, it follows that we may choose a sequence $\{f_q(x)\}$ of functions, each of which satisfies a uniform Lipschitz condition on $[a, b]$ such that

$$\lim_{q\to\infty}\int_a^b |f_q - f|\,dx = 0, \qquad \lim_{q\to\infty}\int_a^b \left|\frac{\partial f_q}{\partial x_i} - D_{x_i}f\right| dx = 0 \qquad (i = 1, \cdots, n).$$

Clearly the functions $g_q(y) = f_q[x(y)]$ each satisfy a uniform Lipschitz condition on $[c, d]$, and it follows from (4) of Lemma 6.2 that

$$\lim_{q\to\infty}\int_c^d |g_q - g|\,dy = 0. \tag{6.4}$$

It also follows from (4) and (7) of Lemma 6.2 that

$$\lim_{q\to\infty}\int_c^d \left|\frac{\partial g_q}{\partial y_k} - G_k(y)\right| dy = 0, \tag{6.5}$$

if we define $G_k(y)$ (summable by (4) of Lemma 6.2, $\partial x_i/\partial y_k$ being bounded) by

$$G_k(y) = \sum_{i=1}^n D_{x_i}f[x(y)]\cdot\frac{\partial x_i}{\partial y_k}. \tag{6.6}$$

Thus, by Theorem 4.2, it follows that $g(y)$ is of class $\mathfrak{P}$ on T. If f is of class $\mathfrak{P}''$ on S, we may choose (on account of Theorem 5.5) the $\{f_q\}$ to converge uniformly on $[a, b]$ to f, in which case, the $\{g_q\}$ converge uniformly to g and g is of class $\mathfrak{P}''$ on T.

From the approximation (6.4) and (6.5), the definition (6.6), and the fact that $G_k(y) = D_{y_k} g$ almost everywhere on $[c, d]$, we see (using (4) of Lemma 6.2) that if $s \subset \Sigma$ and $t \subset [c, d]$, then

$$\int_t D_{y_k}g\,dy = \int_t \sum_{i=1}^n (D_{x_i}f)\,\frac{\partial x_i}{\partial y_k}\,dy = \int_s \sum_{i=1}^n (D_{x_i}f)\cdot\left(\frac{\partial x_i}{\partial y_k}\right)\cdot\left|\frac{dy}{dx}\right| dx.$$

Moreover, if $y_0 \in (c, d)$, $x_0 = x(y_0)$, $\{t\}$ is a regular family at y_0, s corresponds to t, and the $x_i(y)$ all have total differentials at y_0, we see that $\{s\}$ forms a regular family about x_0. Hence, if $(D_{x_i}f)_{x_0}$ exists $(i = 1, \cdots, n)$, we have

$$\lim_{t\to y_0}\frac{1}{m(t)}\int_t (D_{y_k}g)\,dy = \lim_{t\to y_0}\frac{m(s)}{m(t)}\cdot\frac{1}{m(s)}\int_s\left[\sum_{i=1}^n (D_{x_i}f)\cdot\frac{\partial x_i}{\partial y_k}\right]\cdot\left|\frac{dy}{dx}\right| dx$$

and the limit on the right exists and is equal to the right side of (6.3) (using (8) of Lemma 6.2). But the limit on the left was defined to be $(D_{y_k}g)_{y_0}$. Thus the formulas (6.3) are established.

If the transformation is also regular, it is clear from (6) of Lemma 6.2 that

all the partials $\partial x_i/\partial y_j$ and $\partial y_i/\partial x_j$ are uniformly bounded on T and S respectively and the Jacobians dx/dy and dy/dx are uniformly bounded and bounded away from zero on T and S respectively. In this case, it is clear, if we use the formulas (6.3) and also (4) of Lemma 6.2, that $g(y)$ is of class $\mathfrak{P}_\alpha$ ($\mathfrak{P}''_\alpha$) on T if $f(x)$ is of class $\mathfrak{P}_\alpha$ ($\mathfrak{P}''_\alpha$) on S.

It is to be noted that no statement has been made concerning the functions of class $\mathfrak{P}'$ or $\mathfrak{P}'_\alpha$. In this connection, we may point out that if $f(x)$ is of class $\mathfrak{P}$, the spherical h-average function is of class C′ and Lemma 3.2 holds if f_h is replaced by the spherical average. It is therefore clear that a function of class $\mathfrak{P}$ is equivalent to a function of class $\mathfrak{P}'$ which remains of class $\mathfrak{P}'$ under any rotation of axes and hence under any linear transformation. A much stronger statement is contained in Theorem 6.3 below. Before proving this, we prove a theorem considerably stronger than Lemma 3.2. This result is stated in Theorem 6.2.

THEOREM 6.2. *Let $z(x)$ be E.A.C. in x_k on the region G. Let $\Phi(e)$ be a set function defined on G by the formula*

$$\Phi(e) = \int_e \varphi(x)\, dx, \qquad 0 < \gamma \leqq \varphi(x) \leqq \Gamma, \tag{6.7}$$

$\varphi(x)$ being measurable on G. Then there exists a subset Z_k of G_k, the projection of G on the hyperplane $x_k = 0$, Z_k being of $(n-1)$-dimensional measure zero, such that if $\bar{x}'_k$ is not in Z_k we have

$$\lim_{e\to\bar{x}} [\Phi(e)]^{-1} \int_e z(x)\varphi(x)\, dx = \bar{z}(\bar{x}) \tag{6.8}$$

exists for each $\bar{x}_k$ such that $\bar{x} = (\bar{x}'_k, \bar{x}_k)$ is in G, $\{e\}$ being any regular family of sets at $\bar{x}$ (see (6.11) below)*; furthermore, the limit $\bar{z}(\bar{x})$ is independent of the family of sets and the set function $\Phi(e)$ and $\bar{z}(\bar{x}'_k, x_k)$ is A.C. in x_k on any closed segment of the line $x'_k = \bar{x}'_k$ which is interior to G.*

In Lemma 3.2 we have proved the above for the case that $\Phi(e) = m(e)$ ($\varphi \equiv 1$), the sets e in (6.8) being restricted to be hypercubes with center at $\bar{x}$. We let $\bar{z}(x)$ denote the limiting function defined by (6.8) under these conditions; it is L.A.C. in x_k if we define it to be zero, say, where the limit fails to exist.

Now, for each pair (p, q) of integers, $-\infty < p < \infty$, $q \geqq 1$, we define the functions $z^+_{p,q}(x)$ and $z^-_{p,q}(x)$ as follows:

$$z^+_{p,q}(x) = \begin{cases} \bar{z}(x) - p/q, & \bar{z}(x) \geqq p/q, \\ 0, & \bar{z}(x) < p/q; \end{cases} \tag{6.9}$$
$$z^-_{p,q}(x) = \begin{cases} 0, & \bar{z}(x) \geqq p/q, \\ p/q - \bar{z}(x), & \bar{z}(x) < p/q. \end{cases}$$

It is easy to see that all the $z^+_{p,q}(x)$ and $z^-_{p,q}(x)$ are L.A.C. in x_k on G and are non-negative. Hence for each (p, q) there exists a set $Z_{k,p,q}$ of measure zero such that if $\bar{x}'_k$ is not in $Z_{k,p,q}$ we have

$$(6.10)\qquad \begin{aligned} \lim_{h\to 0} (2h)^{-n} \int_{\bar{x}-h}^{\bar{x}+h} z^+_{p,q}(x)\,dx &= z^+_{p,q}(\bar{x}),\\ \lim_{h\to 0} (2h)^{-n} \int_{\bar{x}-h}^{\bar{x}+h} z^-_{p,q}(x)\,dx &= z^-_{p,q}(\bar{x}) \end{aligned}$$

for any $\bar{x}_k$ such that $\bar{x} = (\bar{x}'_k, \bar{x}_k)$ is in G. Let $Z_k = \sum_{p=-\infty}^{\infty} \sum_{q=1}^{\infty} Z_{k,p,q}$ and Z_k is of measure zero and if $\bar{x}'_k$ is not in Z_k, (6.9) holds for every pair (p, q) ($\bar{x}$ being in G as above).

Now, suppose $\bar{x}'_k$ is not in Z_k and $(\bar{x}'_k, \bar{x}_k)$ is in G. Let $\{e\}$ be any regular family of sets at $\bar{x}$. Then, by definition, each e is in a cell $[\bar{x} - h, \bar{x} + h]$ in G such that

$$(6.11)\qquad m[\bar{x} - h, \bar{x} + h] \leqq K\cdot m(e),$$

where K depends only on the family and not on the individual set e, and each such cell with $h > 0$ contains one of the sets e. For each q, choose p_q so that $p_q/q \leqq \bar{z}(\bar{x}) < (p_q + 1)/q$. Then, for each e of $\{e\}$, we have

$$(6.12)\qquad \begin{aligned} 0 &\leqq [\Phi(e)]^{-1} \int_e z^+_{p_q,q}(x)\varphi(x)\,dx \leqq \frac{K\Gamma}{\gamma(2h)^n} \int_{\bar{x}-h}^{\bar{x}+h} z^+_{p_q,q}(x)\,dx,\\ 0 &\leqq [\Phi(e)]^{-1} \int_e z^-_{p_q,q}(x)\varphi(x)\,dx \leqq \frac{K\Gamma}{\gamma(2h)^n} \int_{\bar{x}-h}^{\bar{x}+h} z^-_{p_q,q}(x)\,dx, \end{aligned}$$

using (6.7) and (6.11). Thus, from (6.12) and (6.10) and our choice of p_q, it follows that

$$0 \leqq \liminf_{e\to\bar{x}} [\Phi(e)]^{-1} \int_e z^+_{p_q,q}(x)\varphi(x)\,dx \leqq \limsup_{e\to\bar{x}} [\Phi(e)]^{-1} \int_e z^+_{p_q,q}(x)\varphi(x)\,dx \leqq \frac{K\Gamma}{\gamma q},$$

$$\lim_{e\to\bar{x}} [\Phi(e)]^{-1} \int_e z^-_{p_q,q}(x)\varphi(x)\,dx = 0, \qquad e \in \{e\}.$$

From this and (6.9) we see that, for each q,

$$\frac{p_q}{q} \leqq \liminf_{e\to\bar{x}} [\Phi(e)]^{-1} \int_e z(x)\varphi(x)\,dx \leqq \limsup_{e\to\bar{x}} [\Phi(e)]^{-1} \int_e z(x)\varphi(x)\,dx \leqq \frac{p_q}{q} + \frac{K\Gamma}{\gamma q}.$$

From this and the choice of p_q, it follows immediately that

$$\lim_{e\to\bar{x}} [\Phi(e)]^{-1} \int_e z(x)\varphi(x)\,dx = \bar{z}(\bar{x}),$$

and the theorem is proved.

THEOREM 6.3. *Any function of class* $\mathfrak{P}$ *on a region* G *is equivalent to a function which is of class* $\mathfrak{P}'$ *on* G *and is also of class* $\mathfrak{P}'$ *in any other coördinate system related to the original by a transformation of class* K; *i.e., if* $\bar{z}(x)$ *is such a function and* $x = x(y)$ *is a transformation of class* K *of a region* H *into* G, *and if* $\bar{w}(y) = \bar{z}[x(y)]$, *then* $\bar{w}(y)$ *is of class* $\mathfrak{P}'$ *on* H. *For* $\bar{z}(x)$, *we may take the function defined by* (6.8).

For, let $\bar{z}(x)$ be this function and let $x = x(y)$ be such a transformation. Then if G' and H' are corresponding regions in which $\bar{G}' \subset G$ and $\bar{H}' \subset H$, the above transformation is also regular from H' to G'; and if we define $\Psi(f) = \Phi(e)$ whenever f and e are corresponding sets in H' and G', we see from Lemma 6.2 that regular families $\{f\}$ and $\{e\}$ correspond, and that

$$\Psi(f) = \int_f \varphi[x(y)] \left| \frac{dx}{dy} \right| dy, \qquad 0 < \delta \leqq \varphi[x(y)] \cdot \left| \frac{dx}{dy} \right| \equiv \psi(y) \leqq \Delta,$$

$$\lim_{f \to \bar{y}} [\Psi(f)]^{-1} \int_f \bar{w}(y)\psi(y)\, dy = \lim_{e \to \bar{x}} [\Phi(e)]^{-1} \int_e \bar{z}(x)\varphi(x)\, dx$$

whenever $\bar{x} = x(\bar{y})$. Thus $\bar{w}(y)$ is of class $\mathfrak{P}'$ on each such H' and hence on H, since $\bar{w}(y)$ is certainly of class $\mathfrak{P}$ by Theorem 6.1.

7. **Boundary values for regions of class** K. (See [4], pp. 311–329.) As the functions of class $\mathfrak{P}_\alpha$ and $\mathfrak{P}'_\alpha$ are not continuous, the sense in which these functions take on boundary values is of great importance. In this section we shall study this question for functions of these classes which are defined on certain sufficiently regular regions which we shall now proceed to define. If G is a region and $\bar{G}$ is its closure, neighborhoods on $\bar{G}$ are merely sets open in $\bar{G}$ (i.e., products of $\bar{G}$ with open sets).

DEFINITION 7.1. A region G is said to be *of class* K if $\bar{G}$ may be covered by a finite number $\gamma_1, \cdots, \gamma_N$ of neighborhoods on $\bar{G}$, each of which is the image under a regular transformation $x = x_j(y)$ $(j = 1, \cdots, N)$, of class K, of either the unit hypercube $R_1 : |y_i| < 1$ $(i = 1, \cdots, n)$ or the part (half cube) R_2 of R_1 for which $y_n \leqq 0$, n being the dimensionality. In the latter case we assume that the transformation T_j establishes a 1-1 correspondence between the points of $G^* \cdot \gamma_j$ and the points of R_2 for which $y_n = 0$. If all the transformations above are also of class C′, we say that G is of *class* C′.

It thus appears that if G is of class K, then $\bar{G}$ is a compact manifold (with boundary) of "class K" in the sense of differential geometry ([12], Chapters II and III) (the class K referring, as is customary, to the class of the transformations set up between overlapping coördinate systems), and that if G is of class C′, $\bar{G}$ is a compact manifold (with boundary) of class C′ and G^* consists of a finite number of (regular) manifolds of class C′.

In Lemma 6.2, we observed that if $x = x(y)$ is a regular transformation of class K, then dx/dy is of one sign almost everywhere. If $dx/dy > 0$ almost everywhere, we shall say that the transformation is *positive*. We shall state a

theorem concerning positive transformations which is obvious if the transformation is of class C′ and is easily verified in general. The theorem is as follows:

LEMMA 7.1. *Let $x = x(y)$ be a positive regular transformation of class* K *of a cell T: $c_i < y_i < d_i$ $(i \neq k)$, $c_k < y_k \leqq 0$ into a subset S of the cell $a_i < x_i < b_i$ $(i \neq k)$, $a_k < x_k \leqq 0$ which sets up a* 1-1 *correspondence between the points T^* of T for which $y_k = 0$ and the points S^* of S for which $x_k = 0$. Then the transformation $x_k' = x_k'(y_k', 0)$ is a positive regular transformation of the set T^* (as a set in the y_k'-space) into S^* (as a set in the x_k'-space).*

We next define certain completely additive set functions $x_k'(e)$ and $s(e)$ for certain sets e on the boundary G^* of a region G of class K. Their significance will become clear in the sequel.

LEMMA 7.2. *If G is of class* K, *there exist unique completely additive set functions $x_k'(e)$ and $s(e)$, defined for certain sets e on G^*, which have the following properties*:

(i) *$s(e) \geqq 0$ and $x_k'(e)$ and $V_{x_k'}(e)$ are both defined whenever $s(e)$ is defined and they satisfy the relation $| x_k'(e) | \leqq V_{x_k'}(e) \leqq s(e)$, $V_{x_k'}(e)$ denoting the variation of x_k' over e;*

(ii) *if $s(e) = 0$ and e_k is the projection of e on the hyperplane $x_k = 0$, then e_k is of measure zero;*

(iii) *if $x = x(y)$ is a positive regular transformation, of class* K, *of a cell R: $a_i < x_i < b_i$ $(i \neq k)$, $a_k < y_k \leqq 0$ into a portion γ of $\bar{G}$ as in Definition* 7.1, *and if the set e on $\gamma \cdot G^*$ corresponds to a measurable set E on the part of R where $y_k = 0$, then $x_k'(e)$, $V_{x_k'}(e)$, and $s(e)$ are defined and we have*

$$x_k'(e) = \int_E \frac{\partial x_k'}{\partial y_k'}\, dy_k', \qquad V_{x_k'}(e) = \int_E \left| \frac{\partial x_k'}{\partial y_k'} \right| dy_k', \tag{7.1}$$

$$s(e) = \int_E \left[\sum_{i=1}^{n} \left(\frac{\partial x_i'}{\partial y_k'} \right)^2 \right]^{\frac{1}{2}} dy_k'.$$

Moreover, if e is normal ([11], pp. 83–86) *with respect to $s(e)$, then E is measurable.*

If $X = X(x)$ is a regular transformation of class K *of $\bar{G}$ into $\bar{H}$, then sets measurable with respect to $s(e)$ and $S(E)$ (on H^*) correspond and there exists a constant $K \geqq 1$, depending only on the transformation, such that*

$$K^{-1}s(e) \leqq S(E) \leqq Ks(e),$$

whenever e and E are corresponding measurable sets.

To prove this, let $\gamma_1, \cdots, \gamma_N$ be a covering of $\bar{G}$ as in Definition 7.1. If j is an integer such that γ_j corresponds to the half cube R_2, then $\gamma_j \cdot G^*$ is open on G^*. We may define the completely additive set function $x_{k,j}'(e)$ on $\gamma_j \cdot G^*$ by first remapping γ_j on R_{2k} (R_2 with k replacing n) as above but assuring that $x = x_{k,j}(y)$ is a positive transformation and then defining $x_{k,j}'(e)$ by (7.1).

Now, let $x = x(z)$ be a positive regular transformation as in the statement of the lemma, where $\gamma \cdot G^*$ is contained in $\gamma_j \cdot G^*$ for some j. Then the trans-

formation $y = y_{k,j}(z)$ is positive and regular and of class K so that the transformation $y'_k = y_{k,j}(z'_k, 0)$ is positive, regular, and of class K. Thus, from Lemma 6.2, it follows that measurable sets E and $E_{k,j}$ correspond and that

$$\int_{E_{k,j}} \frac{\partial x'_k}{\partial y'_k} dy'_k = \int_E \frac{\partial x'_k}{\partial z'_k} dz'_k = x'_{k,j}(e),$$

E and $E_{k,j}$ both corresponding to e. It is clear that the set functions $x'_{k,j}(e)$ join up to form completely additive set functions over the whole of G^*. The formulas (7.1) for any representation follow by addition. Statement (i) is now obvious.

To demonstrate (ii) it is sufficient to observe that any set e measurable (i.e., normal; see [11], pp. 83–86) with respect to $s(e)$ on one of the sets $\gamma_j \cdot G^*$ corresponds, under $x = x_{k,j}(y)$, to a measurable set E on $y_k = 0$ and E is carried into e_k by the transformation $x'_k = x_{k,j}(y'_k, 0)$ which satisfies a uniform Lipschitz condition for $|y'_k| < 1$ (but is not necessarily 1-1). Such a transformation clearly carries sets E of measure zero into sets e_k of measure zero; (ii) then follows by addition.

To demonstrate the last statement, we observe that if $\gamma_1, \cdots, \gamma_N$ is a covering of $\bar{G}$ as in Definition 7.1, $X = X(x)$ sets up a corresponding covering $\Gamma_1, \cdots, \Gamma_N$ of $\bar{H}$. If γ_j and Γ_j have corresponding representations on R_{2k}, then corresponding sets e and E on $\gamma_j \cdot G^*$ and $\Gamma_j \cdot H^*$ correspond, by (iii), to the same set f on $y_k = 0$; and if either e or E is measurable (i.e., normal with respect to $s(e)$ or $S(E)$), f is and the other one of e or E is likewise. The last statement follows from the formulas (7.1).

It is clear that $s(e)$ is the hyperarea function on G^*. Also, if G is of class K and also of the form $x'_k \in G_k$, $\varphi_1(x'_k) < x_k < \varphi_2(x'_k)$, then $x'_k(e) = m(e_{2,k}) - m(e_{1,k})$, where e_i is the part of e for which $x_k = \varphi_i(x'_k)$ and $e_{i,k}$ is the projection of e_i on the hyperplane $x_k = 0$ $(i = 1, 2)$.

We now recall the definition of strong convergence suggested in §5.

Definition 7.2. We say that a sequence $\{z_p(x)\}$ *converges strongly in* $\mathfrak{P}_\alpha$ *on* G *to* $z(x)$ if all the $z_p(x)$ and $z(x)$ are of class $\mathfrak{P}_\alpha$ on G and if $\bar{D}_\alpha(z_p - z, G) \to 0$, i.e.,

$$\lim_{p \to \infty} \int_G \left[|z_p - z|^\alpha + \left\{ \sum_{i=1}^n D^2_{x_i}(z_p - z) \right\}^{\frac{1}{2}\alpha} \right] dx = 0.$$

Before proving the first fundamental theorem of this section, we prove the following important lemma:

Lemma 7.3. *Let* $z(x)$ *be of class* $\mathfrak{P}_\alpha$ *on a cell* R: (a, b) *and let* $\bar{z}(x)$ *be the equivalent function defined by* (6.8). *Then*

(i) *for almost every* $\bar{x}'_k$ *in* (a'_k, b'_k), $\bar{z}(\bar{x}'_k, x_k)$ *tends to a limit as* x_k *tends to either* a_k *or* b_k, *and if* $\bar{z}(\bar{x}'_k, a_k)$ *and* $\bar{z}(\bar{x}'_k, b_k)$ *are defined by these limits, then* $\bar{z}(\bar{x}'_k, x_k)$ *is A.C. for* $a_k \leqq x_k \leqq b_k$ *with* $|\partial\bar{z}/\partial x_k|^\alpha$ *summable* $(k = 1, \cdots, n)$;

(ii) $z(x)$ *may be extended to be of class* $\mathfrak{P}$ *over the whole space and of class* $\mathfrak{P}_\alpha$ *over any bounded region;*

(iii) *there exists a sequence* $\{z_p(x)\}$, *each of class* C′ *on* $\bar{R}$, *which converges strongly in* $\mathfrak{P}_\alpha$ *on* R *to* $z(x)$;

(iv) *if* $\{z_p(x)\}$ *is any sequence, each satisfying a uniform Lipschitz condition on* $\bar{R}$, *which converges strongly in* $\mathfrak{P}_\alpha$ *on* R *to* $z(x)$, *then* $z_p(x_k', a_k)$ *and* $z_p(x_k', b_k)$ *converge strongly in* L_α *(in* x_k'*) to* $\bar{z}(x_k', a_k)$ *and* $\bar{z}(x_k', b_k)$, *i.e.*,

$$\lim_{p\to\infty} \int_{a_k'}^{b_k'} [| z_p(x_k', a_k) - \bar{z}(x_k', a_k) |^\alpha + | z_p(x_k', b_k) - \bar{z}(x_k', b_k) |^\alpha] \, dx_k' = 0;$$

(v) *if* $\bar{z}$ *vanishes (almost everywhere) on* R^* *and* $z(x)$ *is defined to be zero on and outside* R^*, *then* z *is of class* $\mathfrak{P}_\alpha$ *over the whole space and the functions* $z_p(x)$ *may be chosen so that each vanishes on and near* R^*, *and hence may also be taken to be of class* $C^{(\infty)}$ *on* $\bar{R}$.

The first statement is obvious and the extension in (ii) may be carried out as follows: (1) define $z(x) = \bar{z}(x)$ on R^* (being zero where $\bar{z}$ is not defined), (2) define $z(x)$ on $[a, 2b - a]$ by repeated application of the formulas $z(x_k', 2b_k - x_k) = z(x_k', x_k)$ $(a_k \leqq x_k \leqq b_k;\ k = 1, \cdots, n)$, and (3) define $z(x_k', x_k + 2b_k - 2a_k) = z(x_k', x_k)$ $(k = 1, \cdots, n)$. (iii) follows from (ii) and Theorem 3.4 and the first part of (v) is obvious.

To prove (iv), let $\{z_q(x)\}$ be any subsequence of $\{z_p(x)\}$. There exists a subsequence $\{z_r(x)\}$ of $\{z_q(x)\}$ such that $z_r(x_k', \bar{x}_k)$ converges strongly in L_α on (a_k', b_k') to $\bar{z}(x_k', \bar{x}_k)$ for almost all $\bar{x}_k$ $(a_k < \bar{x}_k < b_k;\ k = 1, \cdots, n)$. Now fix k and choose $\epsilon > 0$. We may evidently choose $\bar{x}_k$ so near a_k (for example) that

$$(7.2)\qquad \begin{aligned} &\left\{\int_{a_k'}^{b_k'} | \bar{z}(x_k', \bar{x}_k) - \bar{z}(x_k', a_k) |^\alpha dx_k'\right\}^{1/\alpha} \\ &\qquad \leqq (\bar{x}_k - a_k)^{1-1/\alpha} \left\{\int_{a_k'}^{b_k'} \int_{a_k}^{\bar{x}_k} | D_{x_k} z |^\alpha dx\right\}^{1/\alpha} < \tfrac{1}{3}\epsilon, \\ &\left\{\int_{a_k'}^{b_k'} | z_r(x_k', \bar{x}_k) - z_r(x_k', a_k) |^\alpha dx_k'\right\}^{1/\alpha} \\ &\qquad \leqq (\bar{x}_k - a_k)^{1-1/\alpha} \left\{\int_{a_k'}^{b_k'} \int_{a_k}^{\bar{x}_k} | D_{x_k} z_r |^\alpha dx\right\}^{1/\alpha} < \tfrac{1}{3}\epsilon \quad (r = 1, 2, \cdots), \end{aligned}$$

and so that $z_r(x_k', \bar{x}_k)$ converges strongly in L_α on (a_k', b_k') to $\bar{z}(x_k', \bar{x}_k)$. Then we may choose P so that

$$(7.3)\qquad \left\{\int_{a_k'}^{b_k'} | z_r(x_k', \bar{x}_k) - \bar{z}(x_k', \bar{x}_k) |^\alpha dx_k'\right\}^{1/\alpha} < \tfrac{1}{3}\epsilon$$

for all $r > P$. From (7.2), (7.3), and Minkowski's inequality, it follows that

$$\left\{\int_{a_k'}^{b_k'} | z_r(x_k', a_k) - \bar{z}(x_k', a_k) |^\alpha dx_k'\right\}^{1/\alpha} < \epsilon$$

for all $r > P$. Thus $\{z_r(x_k', a_k)\}$ converges strongly in L_α on (a_k', b_k') to $\bar{z}(x_k', a_k)$. From a well-known theorem[3] on mean convergence, it follows that the whole sequence $\{z_p(x_k', a_k)\}$ tends strongly in L_α on (a_k', b_k') to $\bar{z}(x_k', a_k)$.

To prove the last part of (v), we define $z(x) = 0$, for x not interior to R and define

$$\bar{z}_p(x) = z\left[\frac{a+b}{2} + \frac{p+1}{p}\cdot\left(x - \frac{a+b}{2}\right)\right].$$

It is clear that each $\bar{z}_p(x)$ is of class $\mathfrak{P}_\alpha$ over the whole space and zero outside the cell $(\frac{1}{2}(a + b) - p/(p + 1)\cdot\frac{1}{2}(b - a), \frac{1}{2}(a + b) + p/(p + 1)\cdot\frac{1}{2}(b - a))$ and that $\bar{z}_p(x)$ tends strongly in $\mathfrak{P}_\alpha$ on R to $z(x)$. For each p, we may choose h_p so small that the h_p-average function $\bar{z}_{p,h_p}$ of $\bar{z}_p$ is zero on and near R^* and so that $\bar{D}_\alpha(\bar{z}_{p,h_p} - \bar{z}_p, R) < 1/p$. Evidently the sequence $\{z_p(x)\} = \{\bar{z}_{p,h_p}(x)\}$ satisfies the desired conditions.

The following lemma is obvious but useful in the sequel.

LEMMA 7.4. *Let $z(x)$ be of class $\mathfrak{P}$ ($\mathfrak{P}'$, $\mathfrak{P}''$, $\mathfrak{P}_\alpha$, $\mathfrak{P}_\alpha'$, $\mathfrak{P}_\alpha''$) and let $h(x)$ satisfy a uniform Lipschitz condition on the bounded region G. Then $h(x)\cdot z(x)$ is of class $\mathfrak{P}$ ($\mathfrak{P}'$, $\mathfrak{P}''$, $\mathfrak{P}_\alpha$, $\mathfrak{P}_\alpha'$, $\mathfrak{P}_\alpha''$, respectively) on G and its generalized derivatives are given by the usual formulas.*

As the following procedure will be used frequently in the sequel, we shall describe it here and refer to it later by name. It follows easily, from the Heine-Borel theorem and Definition 7.1, that if G is of class K, there exist sets $\Gamma_1, \cdots, \Gamma_N$ and $\gamma_1, \cdots, \gamma_N$ such that (1) each Γ_i is the image, under a regular transformation T_i: $x = x_i(y)$ of class K, of either R_1 or R_2 with restrictions as in Definition 7.1; (2) T_i establishes a correspondence between γ_i and the part of R_1 or R_2, respectively, for which $|y_i| < \frac{1}{2}$ $(i = 1, \cdots, n)$; and (3) $\bar{G}$ is covered by the sets $\gamma_1, \cdots, \gamma_N$.

DEFINITION 7.3. We call such a covering of G a *canonical covering*, and designate it by (Γ, γ, N, T).

Suppose, now, that G is of class K and that (Γ, γ, N, T) is a canonical covering of $\bar{G}$. In R_1, define $h(y)$ as follows:

$$h(y) = \begin{cases} 1, & d(y) \geqq \frac{1}{2}, \\ 1 + 4[d(y) - \frac{1}{2}], & \frac{1}{2} \geqq d(y) \geqq \frac{1}{4}, \\ 0, & \frac{1}{4} \geqq d(y), \end{cases}$$

where $d(y)$ is the distance of y from R_1^*. For each i $(1 \leqq i \leqq N)$, we then define $h_i(x)$ on Γ_i as the transform of $h(y)$ under T_i and define it to be zero elsewhere in $\bar{G}$. Evidently each $h_i(x)$ then satisfies a uniform Lipschitz condi-

[3] The reader can verify the following theorem: *Let x_0 and x_n $(n = 1, 2, \cdots)$ be elements of a metric space E and suppose that each subsequence of $\{x_n\}$ contains a further subsequence which converges to x_0; then the whole sequence $\{x_n\}$ tends to x_0.*

tion on $\bar{G}$. Now, suppose $z(x)$ is of class $\mathfrak{P}_\alpha$ on G. We define functions $z_1(x), \cdots, z_N(x)$ as follows:

$$z_1(x) = h_1(x)z(x), \qquad z_i(x) = h_i(x)\cdot\left[z(x) - \sum_{j=1}^{i-1} z_j(x)\right] \qquad (i = 2, \cdots, N).$$

Then each $z_i(x)$ is of class $\mathfrak{P}_\alpha$ on G and is zero on and near $\overline{G - \Gamma_i}$ and also

$$z(x) = \sum_{i=1}^{N} z_i(x). \tag{7.4}$$

That (7.4) holds is easily seen, since $z_1(x) = z(x)$ on γ_1 so that $z - z_1 = 0$ on γ_1. Also $z_2(x) = z(x) - z_1(x)$ on γ_2 so $z(x) - z_1(x) - z_2(x)$ is zero on $\gamma_1 + \gamma_2$. In general $z(x) - z_1(x) - \cdots - z_i(x)$ is zero on $\gamma_1 + \cdots + \gamma_i$.

Definition 7.4. The particular representation of $z(x)$ given by (7.4) is called *the canonical resolution of $z(x)$ determined* by (Γ, γ, N, T).

We are now in a position to prove the first fundamental theorem of the section.

Theorem 7.1. *Let $z(x)$ be of class $\mathfrak{P}_\alpha$ on a region G of class* K. *Then*

(i) *there exists a sequence $\{z_p(x)\}$ of functions, each satisfying a uniform Lipschitz condition on $\bar{G}$, which converges strongly in $\mathfrak{P}_\alpha$ to $z(x)$ on G;*

(ii) *there exists a function $\varphi(x)$ of class L_α on G^* (with respect to $s(e)$) to which any sequence $\{z_p(x)\}$ in* (i) *converges strongly in L_α on G^*;*

(iii) *if $x' = x'(x)$ is a regular transformation, of class* K, *of $\bar{G}$ into $\bar{G}'$ and if $z'(x') = z[x(x')]$ and $\varphi'(x') = \varphi[x(x')]$, then φ' bears the same relation to z' on G'^* as φ does to z on G^*;*

(iv) *if $\varphi(x) = 0$ (almost everywhere) on G^*, then the functions $z_p(x)$ may be chosen so that each vanishes on and near G^*, and hence if $z(x)$ is defined to be zero for x not in G, then $z(x)$, as extended, is of class $\mathfrak{P}_\alpha$ over the whole space.*

To prove (i), let (Γ, γ, N, T) be a canonical covering of $\bar{G}$, let $z(x) = \sum_{i=1}^{N} z_i(x)$ be the corresponding canonical resolution of $z(x)$, and let $w_i(y)$ be the transform of $z_i(x)$ under T_i $(i = 1, \cdots, N)$. In case $w_i(y)$ is defined only in R_2, extend it to the whole R_1 by the equation $w_i(y_n', y_n) = w_i(y_n', -y_n)$. Then each $w_i(y)$ is of class $\mathfrak{P}_\alpha$ on R_1 and vanishes near R_1^* and hence $\bar{w}_i(y) = 0$ on R_1^*. From Lemma 7.3, it follows that there exists a sequence $\{w_{i,p}(x)\}$, for each i, each satisfying a uniform Lipschitz condition on $\bar{R}_1$ and vanishing on and near R_1^*, which converges strongly in $\mathfrak{P}_\alpha$ on R_1 to $w_i(y)$ $(i = 1, \cdots, N)$. If we define $z_{i,p}(x)$ on Γ_i as the transform, under T_i, of $w_{i,p}(y)$, defining $z_{i,p}(x) = 0$ elsewhere on $\bar{G}$, and if we then define $z_p(x) = \sum_{i=1}^{N} z_{i,p}(x)$, we see that $\{z_p(x)\}$ is a sequence of the desired type.

Now, let us define $\varphi_i(x)$ on $\Gamma_i \cdot G^*$ as the transform under T_i of $\bar{w}_i(y_n', 0)$ for those i for which $\Gamma_i \cdot G^* \neq 0$ and corresponds under T_i to the points of R_2 for which $y_n = 0$; let us define $\varphi_i(x) = 0$ elsewhere on G^*; and let us define $\varphi(x) =$

$\sum_{i=1}^{N} \varphi_i(x)$. Let $\{z_p(x)\}$ be any sequence as in (i); for each p, let $z_p(x) = \sum_{i=1}^{N} z_{i,p}(x)$ be the canonical resolution of $z_p(x)$ determined by (Γ, γ, N, T), and let $w_{i,p}(y)$ be the transform under T_i of $z_{i,p}(x)$ (as defined on Γ_i). Then, from Lemma 7.3, it follows that $w_{i,p}(y_n', 0)$ converges strongly in L_α to $\bar{w}_i(y_n', 0)$. From Lemma 7.2 and the fact that each $z_{i,p}(x)$ is zero on and near $\overline{G - \Gamma_i}$, it follows that $z_{i,p}(x)$ converges strongly in L_α on G^* to $\varphi_i(x)$ $(i = 1, \cdots, N)$. This demonstrates (ii).

The third statement is now obvious by virtue of Lemma 7.2. Moreover, if $\varphi(x) = 0$ on G^*, it is now clear that each $\varphi_i(x)$ is also zero on G^*, and hence each $\bar{w}_i(y)$ is zero on R_1^* or R_2^* as the case may be. From Lemma 7.3, we may find sequences $\{w_{i,p}(y)\}$, each function satisfying a uniform Lipschitz condition on $\bar{R}_1$ or $\bar{R}_2$ and zero near R_1^* or R_2^* (respectively), which converge strongly in $\mathfrak{P}_\alpha$ to $\bar{w}_i(y)$ $(i = 1, \cdots, N)$. The corresponding functions $z_{i,p}(x)$ (zero except on Γ_i) and hence the functions $z_p(x) = \sum_{i=1}^{N} z_{i,p}(x)$ satisfy the conditions in (iv).

Definition 7.5. Let $z(x)$ be of class $\mathfrak{P}_\alpha$ on G and let $\varphi(x)$ be the function of Theorem 7.1 defined on G^*. We say that $z(x)$ *takes on the boundary values* $\varphi(x)$ *on* G^*.

The following two theorems give further interpretations of this boundary value function $\varphi(x)$. The first is an immediate consequence of Theorem 6.3 and of Theorem 7.1 and its proof.

Theorem 7.2. *Let $z(x)$ be of class $\mathfrak{P}_\alpha$ on the region G of class* K, *let $\varphi(x)$ be its boundary values on G^*, and let $\bar{z}(x)$ be the function defined by* (6.8). *Let γ be a portion of $\bar{G}$ which is the image, under the regular transformation T: $x = x(y)$ of class* K, *of the half cube R_2 in which T sets up a correspondence between the points of $\gamma \cdot G^*$ and the points of R_2 for which $y_n = 0$. Let $\bar{w}(y) = \bar{z}[x(y)]$ and $\psi(y_n') = \varphi[x(y_n', 0)]$.*

Then $\bar{w}(y_n', 0)$ (as defined in (i) *of Lemma* 7.3) *$= \psi(y_n')$ for almost all y_n'.*

Theorem 7.3. *Let $z(x)$ be of class $\mathfrak{P}_\alpha$ on a region G of class* K, *let $\varphi(x)$ be its boundary values on G^*, and let $\bar{z}(x)$ be the equivalent function of* (6.8). *Let β be any direction, Π_β be the $(n - 1)$-dimensional hyperplane through the origin and perpendicular to β, and G_β be the projection of G on Π_β.*

Then, if l is any line in the direction β whose intersection with Π_β does not lie in a certain subset Z_β of G_β which is of $(n - 1)$-dimensional measure zero, $\bar{z}(x)$ is A.C. on any segment (x_1^*, x_2^*) of l in G, where x_1^* and x_2^* are on G^* and $\bar{z}(x)$ tends to $\varphi(x_i^*)$ as x tends along this segment to x_i^* $(i = 1, 2)$.

On account of Theorems 6.3 and 7.1, we may take β as the direction of the x_1-axis. Let $\{z_p(x)\}$ be a sequence of functions as in (i) of Theorem 7.1. From a well-known theorem on mean convergence it follows that a subsequence $\{z_q(x)\}$ exists such that $\{z_q(x)\}$ tends to $\varphi(x)$ for almost every (with respect to $s(e)$) x on G^* and such that

$$\lim_{q \to \infty} \int_{l_1} \left| \frac{\partial z_q}{\partial x_1} - \frac{\partial \bar{z}}{\partial x_1} \right|^\alpha dx_1 = 0, \tag{7.5}$$

l_1 being the product of l with G, this convergence holding for all lines l: $x_1' = \bar{x}_1'$ for which $\bar{x}_1'$ does not lie in a set $Z_{1,1}$ of measure zero. From Lemma 7.2, it follows that the projection of the set e, of points x of G^* where $\{z_q(x)\}$ fails to converge to $\varphi(x)$, on $x_1 = 0$ is a set $Z_{1,2}$ of $(n-1)$-dimensional measure zero.

Now, let $Z_1 = Z_{1,1} + Z_{1,2}$; Z_1 is of $(n-1)$-dimensional measure zero. Let $\bar{x}_1'$ be in G (the projection of G on $x_1 = 0$) but not in Z_1, and suppose the points $(\bar{x}_1', x_1^*)$ and $(\bar{x}_1', x_2^*)$ are on G^* and that all the points $(\bar{x}_1', x_1)$ with $x_1^* < x_1 < x_2^*$ are in G. From (7.5), it follows that the functions $z_q(\bar{x}_1', x_1)$ converge uniformly on $x_1^* \leqq x_1 \leqq x_2^*$ to some continuous function which coincides with $\bar{z}(\bar{x}_1', x_1)$ on the open segment, and, since $\bar{x}_1'$ is not in $Z_{1,2}$, $z_q(\bar{x}_1', x_i^*)$ converges to $\varphi(\bar{x}_1', x_i^*)$ $(i = 1, 2)$. This proves the theorem.

We now prove an exceedingly important "substitution theorem". This result is stated in Theorem 7.4.

THEOREM 7.4. *Let $z(x)$ be of class $\mathfrak{P}_\alpha$ $(\alpha \geqq 1)$ on a region G. Let H be a subregion of G with $\bar{H} \subset G$ and suppose that H is of class* K. *Let $u(x)$ be of class $\mathfrak{P}_\alpha$ in H with $D_\alpha(u, H)$ finite, and suppose that $u(x)$ and $z(x)$ have the same (essentially) boundary values (with respect to H) on H^*. Let $w(x)$ coincide with $u(x)$ on H and with $z(x)$ on $G - H$. Then $w(x)$ is of class $\mathfrak{P}_\alpha$ on G.*

Proof. Evidently the function $v(x)$, defined on H by $v(x) = u(x) - z(x)$, is of class $\mathfrak{P}_\alpha$ on H and vanishes on H^*. Now, by Theorem 7.1, the function $V(x)$ which coincides with $v(x)$ on H and is zero on $G - H$ is of class $\mathfrak{P}_\alpha$ over G. Then we see that $w(x) = z(x) + V(x)$ on G and hence $w(x)$ is of class $\mathfrak{P}_\alpha$ on G.

In this section, we have seen that any function of class $\mathfrak{P}_\alpha$ $(\alpha \geqq 1)$ on a region G of class K tends in the mean of order α to boundary values of class L_α. Clearly, if G is the unit sphere in which spherical coördinates (P, r) are chosen, P being on G^*, and if $u(r, P)$ is of class $\mathfrak{P}_\alpha$ on G, then ($s(e)$ being the function of Lemma 7.2)

$$\lim_{r \to 1} \int_{G^*} | \bar{u}(r, P) - \bar{u}(1, P) |^\alpha ds(e) = 0, \tag{7.6}$$

$\bar{u}(r, P)$ being the usual function. Of course (7.6) may hold without u being of class $\mathfrak{P}_\alpha$ on G. In fact, if $u(1, \theta)$ is summable with its square for $0 \leqq \theta \leqq 2\pi$, then (7.6) holds with $\alpha = 2$ if $\bar{u}(r, \theta)$ is the harmonic function given in terms of $\bar{u}(1, \theta)$ by Poisson's integral. However, the Dirichlet integral of this harmonic function need not be finite as

$$D_2[u, C(0, 1)] = \pi \sum_{n=1}^{\infty} n(a_n^2 + b_n^2), \qquad \int_0^{2\pi} [\bar{u}(1, \theta)]^2 d\theta = \pi \left[\tfrac{1}{2} a_0^2 + \sum_{n=1}^{\infty} (a_n^2 + b_n^2) \right],$$

the a_n and b_n being the Fourier coefficients for $u(1, \theta)$.

8. **Weak convergence; Rellich's theorem.** In this section, we consider families and sequences of functions of class $\mathfrak{P}_\alpha$ and of class $\mathfrak{P}$. It is assumed throughout the section that $\alpha \geqq 1$. Weak convergence (as contrasted with strong con-

vergence or convergence with respect to the norm suggested in §7) in $\mathfrak{P}_\alpha$ on a region G is defined and it is shown that any family of functions $\{z(x)\}$ with $\bar{D}_\alpha(z, G)$ uniformly bounded is compact with respect to the above weak convergence except that, for $\alpha = 1$, we must impose a condition of "uniform absolute continuity" on the functions of the family. If G is of class K, it is shown that the weak convergence of $z_p(x)$ to $z(x)$ in $\mathfrak{P}_\alpha$ on G implies the strong convergence in L_α of z_p to z on G and the strong convergence in L_α of the boundary values φ_p of z_p on G^* to the boundary values φ of z. Various sufficient conditions for the compactness in $\mathfrak{P}_\alpha$ of families of functions are developed.

We begin by recalling the definition of weak convergence of elements in a Banach space ([1], Chapters IV, IX).

DEFINITION 8.1. Let X_p and X be elements of a Banach space E. We say that X_p tends weakly to X in E if $F(X_p) \to F(X)$ for each linear functional ([1], p. 23) $F(X)$ defined on E.

It is well known that the spaces L_α ($\alpha \geqq 1$) ([1], Chapter IV) of functions f with $|f|^\alpha$ summable on a region G and the space M of functions measurable and essentially bounded on G are Banach spaces if the norms are defined in the usual way, and we have seen in §5 that the spaces $\mathfrak{P}_\alpha$ ($\alpha \geqq 1$) are also Banach spaces. The following ([1], Chapter IV) is a well-known criterion for weak convergence in the spaces L_α.

LEMMA 8.1. *The most general linear functional $F(f)$ defined on L_α is of the form*

$$F(f) = \int_G A(x)f(x)\,dx,$$

where $A(x)$ is essentially bounded (i.e., $A(x) \in M$) if $\alpha = 1$ and $A(x) \in L_\beta$ with $\alpha^{-1} + \beta^{-1} = 1$ if $\alpha > 1$.

If G is bounded, necessary and sufficient conditions that $f_p(x)$ converge weakly in L_α on G to $f(x)$ are

$$\text{(i)} \qquad \int_G |f_p(x)|^\alpha\,dx \leqq N,$$

$$\text{(ii)} \qquad \lim_{p\to\infty} \int_R f_p(x)\,dx = \int_R f(x)\,dx$$

for each cell R in G, N being independent of p; in the case $\alpha = 1$, we must replace (i) *by the condition that the set functions*

$$\text{(8.1)} \qquad \varphi_p(e) = \int_e f_p(x)\,dx$$

be uniformly absolutely continuous (this is implied by (i) if $\alpha > 1$).

The following facts concerning weak convergence are well known ([1], Chapter IX) or easily deducible from known results.

LEMMA 8.2. *Suppose that* $\{f_p(x)\}$ *converges weakly in* L_α *to* $f(x)$ *on the bounded region* G. *Then*

(1) *the convergence holds on any subregion of* G;

(2) $f(x)$ *is unique (except for a possible additive null function)*;

(3) *we have*

$$\lim_{p\to\infty}\int_e A(x)f_p(x)\,dx = \int_e A(x)f(x)\,dx$$

for any (bounded) measurable subset e *of* G, $A(x)$ *being in* L_β $(\alpha^{-1}+\beta^{-1}=1)$ *or* M *on* G *according as* $\alpha > 1$ *or* $\alpha = 1$; *and*

(4) *we have*

$$\int_G |f(x)|^\alpha\,dx \leqq \liminf_{p\to\infty}\int_G |f_p(x)|^\alpha\,dx.$$

Finally, if $\{f_p(x)\}$ *is any sequence of functions in* L_α *which satisfies* (i) *of Lemma* 8.1 *and in which the set functions of* (8.1) *are uniformly absolutely continuous, then a subsequence of* $\{f_p(x)\}$ *converges weakly in* L_α *on* G *to some function* $f(x)$ *in* L_α.

We now prove the analogues of Lemmas 8.1 and 8.2 for the spaces $\mathfrak{P}_\alpha$.

THEOREM 8.1. *The most general linear functional* $F(z)$ *defined for all* z *of class* $\mathfrak{P}_\alpha$ *on a bounded region* G *is of the form*

$$F(z) = \int_G\left[A(x)z(x) + \sum_{i=1}^n A_i(x)D_{x_i}z\right]dx,$$

where $A(x)$ *and* $A_i(x)$ *are all in* M *if* $\alpha = 1$ *and are all in* L_β *with* $\alpha^{-1}+\beta^{-1}=1$ *if* $\alpha > 1$.

A necessary and sufficient condition that $z_p(x)$ *converge weakly in* $\mathfrak{P}_\alpha$ *on* G *to the function* z *is that* z_p *and* z *are all in* $\mathfrak{P}_\alpha$ *and that* z_p *and* $D_{x_i}z_p$ *tend weakly in* L_α *on* G *to* z *and* $D_{x_i}z$, *respectively* $(i = 1, \cdots, n)$.

To prove this, consider the Banach space E of sets $\varphi = (\varphi_0, \varphi_1, \cdots, \varphi_n)$ of functions, each in L_α on G with

$$\|\varphi\|^\alpha = \int_G\left\{|\varphi_0|^\alpha + \left(\sum_{i=1}^n \varphi_i^2\right)^{\frac{1}{2}\alpha}\right\}dx.$$

Evidently the $n+1$ functions $z, D_{x_1}z, \cdots, D_{x_n}z$ form a certain linear subspace H of E with the same norm, since we have defined (in §7)

$$\|z\|^\alpha = \int_G\left\{|z|^\alpha + \left[\sum_{i=1}^n D_{x_i}^2 z\right]^{\frac{1}{2}\alpha}\right\}dx \equiv \bar{D}_\alpha(z, G).$$

Now, if we are given $F(z)$, we may extend it ([1], Chapter II, Theorem 1) to the whole space E. Now, by holding all the $\varphi_i = 0$ but one, letting i run from 0 to n, and using Lemma 8.1 and the linearity of F, we see that $F(z)$ has the form indicated on E and hence on H.

It is clear in the above theorem that the $A(x)$ and $A_i(x)$ are not uniquely determined by the functional $F(z)$. The following theorem is an immediate consequence of Theorem 8.1 and Lemmas 8.1 and 8.2.

THEOREM 8.2. *Suppose $\{z_p(x)\}$ converges weakly to $z(x)$ in $\mathfrak{P}_\alpha$ on G. Then*
(1) *the convergence holds on any subregion of G;*
(2) *$z(x)$ is unique (except for a possible additive null function);*
(3) *we have*

$$\lim_{p\to\infty}\int_G\left[A(x)z_p(x)+\sum_{i=1}^n A_i(x)D_{x_i}z_p\right]dx=\int_G\left[A(x)z(x)+\sum_{i=1}^n A_i(x)D_{x_i}z\right]dx$$

if all the functions $A(x)$ and $A_i(x)$ are in M if $\alpha = 1$ or in L_β with $\alpha^{-1}+\beta^{-1}=1$ if $\alpha > 1$;
(4) *we have*

$$\int_G |z(x)|^\alpha dx \leqq \liminf_{p\to\infty}\int_G |z_p(x)|^\alpha dx, \qquad \int_G |D_{x_i}z|^\alpha dx \leqq \liminf_{p\to\infty}\int_G |D_{x_i}z_p|^\alpha dx,$$

$$\int_G\left[\sum_{i=1}^n D_{x_i}^2 z\right]^{\frac{1}{2}\alpha}dx \equiv D_\alpha(z, G) \leqq \liminf_{p\to\infty} D_\alpha(z_p, G) \qquad (i = 1, \cdots, n);$$

and
(5) *if $x = x(y)$ is a regular transformation of class* K *of a region H in the y-space into G, and if $w_p(y) = z_p[x(y)]$, $w(y) = z[x(y)]$, then the sequence $\{w_p(y)\}$ converges weakly in $\mathfrak{P}_\alpha$ on H to $w(y)$.*

Before proceeding to a further discussion of weak convergence, we shall prove an important inequality which has been proved for $n = 2$, $\alpha = 2$ by K. Friedrichs ([6], pp. 230–232) for somewhat more general regions.

THEOREM 8.3. *Let $z(x)$ be of class $\mathfrak{P}_\alpha$ on a hypercube R: $[a, b]$ of side h. Then*

(i) $$\int_a^b\int_a^b |z(x)-z(\xi)|^\alpha dx\,d\xi \leqq [(2n-1)\cdot h]^\alpha\cdot h^n\cdot D_\alpha(z, R);$$

(ii) $$\int_a^b |z(x)-z_R|^\alpha dx \leqq [(2n-1)\cdot h]^\alpha D_\alpha(z, R),$$

where z_R denotes the average of z over R.

As we can approximate to $z(x)$ on R *strongly* in $\mathfrak{P}_\alpha$ by functions of class C′ on $\bar{R}$, it follows that it is sufficient to prove the inequality for such functions. The proof for $n = 2$ is typical, so we shall take $x_1 = x$, $x_2 = y$, $\xi_1 = \xi$, $\xi_2 = \eta$ and R: $(a, b; a + h, b + h)$.

Now, there exists a value $\bar{y}$ $(b \leqq \bar{y} \leqq b + h)$ such that

(8.2) $$\int_a^{a+h}\left|\frac{\partial z}{\partial x}\right|^\alpha dx \leqq h^{-1}D_\alpha(z, R), \qquad y = \bar{y}.$$

Then

$$z(x, y) - z(\xi, \eta) = \int_\eta^{\bar{y}} \frac{\partial z(\xi, t)}{\partial y}\, dt + \int_\xi^x \frac{\partial z(s, \bar{y})}{\partial x}\, ds + \int_{\bar{y}}^y \frac{\partial z(x, t)}{\partial y}\, dt.$$

Hence

$$\left[\int_a^{a+h}\int_b^{b+h}\int_a^{a+h}\int_b^{b+h} |z(x, y) - z(\xi, \eta)|^\alpha\, dx\, dy\, d\xi\, d\eta\right]^{1/\alpha}$$

$$\leqq \left[\iint_R \iint_R \left|\int_\eta^{\bar{y}} \frac{\partial z(\xi, t)}{\partial y}\, dt\right|^\alpha dx\, dy\, d\xi\, d\eta\right]^{1/\alpha}$$

$$+ \left[\iint_R \iint_R \left|\int_\xi^x \frac{\partial z(s, \bar{y})}{\partial x}\, ds\right|^\alpha dx\, dy\, d\xi\, d\eta\right]^{1/\alpha}$$

$$+ \left[\iint_R \iint_R \left|\int_{\bar{y}}^y \frac{\partial z(x, t)}{\partial y}\, dt\right|^\alpha dx\, dy\, d\xi\, d\eta\right]^{1/\alpha} \leqq 3[h^{\alpha+2} D_\alpha(z, R)]^{1/\alpha},$$

and (i) follows immediately; the Hölder inequality was applied to the interior integrals; (ii) is an immediate consequence of (i) as is easily seen by using the definition of z_R and the Hölder inequality.

Before proceeding to the proof of Rellich's theorem [9] in the general case, we shall prove it for a cell.

LEMMA 8.3. *Let the functions $z_p(x)$ $(p = 1, 2, \cdots)$ be of class $\mathfrak{P}_\alpha$ on a cell R: (a, b) with $\bar{D}_\alpha(z_p, R) \leqq M$, M being independent of p. Then there exists a subsequence $\{z_q(x)\}$ which converges strongly in L_α to some function $z(x)$.*

We may extend each z_p by successive reflections to be of class $\mathfrak{P}$ over the whole space. It is clear that the hypotheses above will be fulfilled on some hypercube which contains R. Hence we may assume that R is a hypercube of side K, say.

For each r, we divide R into 2^{nr} congruent hypercubes $R_{r,j}$ $(j = 1, \cdots, 2^{nr})$ each of side $K \cdot 2^{-r}$ by means of hyperplanes perpendicular to the axes. For each p and r, we define the step function $z_{p,r}(x)$ to be equal in $R_{r,j}$ to the average of z_p over $R_{r,j}$. From the above theorem, it follows that, for each p, we have

$$(8.3)\quad \begin{aligned} &\int_R |z_p(x) - z_{p,r}(x)|^\alpha dx \leqq [(2n-1)\cdot K \cdot 2^{-r}]^\alpha \cdot M, \\ &\int_R |z_{p,r}(x)|^\alpha dx = \sum_{j=1}^{2^{nr}} \int_{R_{r,j}} \left| [m(R_{r,j})]^{-1} \int_{R_{r,j}} z_p(x)\, dx \right|^\alpha dx \leqq \int_R |z_p(x)|^\alpha dx \leqq M. \end{aligned}$$

Now, for each fixed r, all the functions $z_{p,r}(x)$ are uniformly bounded and each constant on each $R_{r,j}$. Hence a subsequence $\{p_1\}$ of $\{p\}$ may be chosen so that the functions $z_{p_1,1}(x)$ converge to a function $z_1(x)$ which is a step function of the above type. We may find subsequences $\{p_r\}$ in succession, each being a subsequence of the preceding so that the functions $z_{p_r,r}(x)$ converge to a

function $\bar{z}_r(x)$. If we let $\{q\}$ be the diagonal sequence, we see that $\{z_{q,r}(x)\} \to \bar{z}_r(x)$ for each r.

Now, from (8.3) we see that

$$\left[\int_R |z_{q,r_1}(x) - z_{q,r_2}(x)|^\alpha dx\right]^{1/\alpha} \leqq 2M^{1/\alpha}\cdot(2n-1)\cdot K\cdot 2^{-\bar{r}}, \tag{8.4}$$

where $\bar{r}$ is the smaller of r_1 and r_2. From our above limiting process, we see that

$$\left[\int_R |\bar{z}_{r_1}(x) - \bar{z}_{r_2}(x)|^\alpha dx\right]^{1/\alpha} \leqq 2M^{1/\alpha}\cdot(2n-1)\cdot K\cdot 2^{-\bar{r}}. \tag{8.5}$$

Hence the functions $\bar{z}_r(x)$ converge in L_α to some function $z(x)$. Using (8.3)' (8.4), and (8.5), we easily see that the subsequence $\{z_q(x)\}$ converges strongly in L_α to this function $z(x)$.

We are now able to prove the compactness of bounded subsets of the space $\mathfrak{P}_\alpha$ with respect to weak convergence if G is bounded. This fact is stated in the following theorem:

THEOREM 8.4. *Let $\{z_p(x)\}$ be a sequence of functions of class $\mathfrak{P}_\alpha$ on G with $\bar{D}_\alpha(z_p, G)$ uniformly bounded, G being bounded. Then a subsequence $\{z_q(x)\}$ converges weakly in $\mathfrak{P}_\alpha$ on G to some function $z(x)$ (of class $\mathfrak{P}_\alpha$), $\alpha > 1$. If $\alpha = 1$, the theorem holds if the functions $z_p(x)$ are uniformly of class $\mathfrak{P}_1$, i.e., if the set functions $\varphi_p(e)$ and $\psi_{i,p}(e)$ are uniformly A.C., where*

$$\varphi_p(e) = \int_e z_p(x)\,dx, \qquad \psi_{i,p}(e) = \int_e D_{x_i} z_p\,dx.$$

It is clear from Lemma 8.2 that we may choose a subsequence $\{z_q(x)\}$ such that $\{z_q(x)\}$ converges weakly in L_α to some function $z(x)$ and $\{D_{x_i} z_q\}$ converges weakly in L_α on G to some function $v_i(x)$ $(i = 1, \cdots, n)$.

Now, let $[a, b]$ be any cell of G. From Lemma 8.3 and the uniqueness of $z(x)$, it follows that a subsequence $\{z_r(x)\}$ tends strongly in L_α to $z(x)$. Hence if $\{z_s(x)\}$ is a properly chosen subsequence, we see that, on almost every cell R of $[a, b]$,

$$\lim_{s\to\infty} \int_{R^*} |z_s(x) - z(x)|^\alpha ds(e) = 0$$

($s(e)$ and $x_k'(e)$ being the set functions of Lemma 7.2 for R^*). Thus, for almost every R, we have

$$\int_{R^*} z(x)\,dx_k'(e) = \lim_{s\to\infty} \int_{R^*} z_s(x)\,dx_k'(e) = \lim_{s\to\infty} \int_R D_{x_k} z_s\,dx = \int_R v_k(x)\,dx,$$

so that $z(x)$ is of class $\mathfrak{P}_\alpha$ on (a, b) and the functions $v_k(x)$ are equivalent respectively to $D_{x_k} z$ $(k = 1, \cdots, n)$. Hence $\{z_q(x)\}$ converges weakly in our sense to $z(x)$.

We now prove an important theorem which, together with Theorem 8.4, includes Rellich's theorem for regions of class K.

THEOREM 8.5. *If $\{z_p(x)\}$ converges weakly in $\mathfrak{P}_\alpha$ on a bounded region G to the function $z(x)$, then $\{z_p(x)\}$ tends strongly in L_α to $z(x)$ on each closed cell interior to G. If G is of class* K, *then $\{z_p(x)\}$ tends strongly in L_α to $z(x)$ on G, and the boundary values of z_p tend strongly in L_α on G^* to those for z.*

To prove the first statement, let (a, b) be any cell in G and let $\{z_q(x)\}$ be any subsequence of $\{z_p(x)\}$. By Lemma 8.3, a subsequence $\{z_r(x)\}$ of $\{z_q(x)\}$ exists which converges strongly in L_α on (a, b) to some function which must be $z(x)$ since $z(x)$ is unique by Lemma 8.2. Thus, by a well-known theorem on mean convergence (see footnote 3), it follows that the whole sequence $\{z_p(x)\}$ tends strongly in L_α to $z(x)$ on (a, b).

Now, suppose that G is of class K. Let (Γ, γ, N, T) be a canonical covering of $\bar{G}$, let $z_p(x) = \sum_{i=1}^{N} z_{i,p}(x)$ and $z(x) = \sum_{i=1}^{N} z_i(x)$ be the corresponding canonical resolutions of $z_p(x)$ and $z(x)$, respectively, and let $w_{i,p}(y)$ and $w_i(y)$ be the transforms, under T_i, of $z_{i,p}(x)$ and $z_i(x)$ respectively as defined on Γ_i; let $\bar{z}_{i,p}$, etc., be the corresponding functions (6.8). From Lemma 7.4 and Theorem 8.2 it follows that $z_{i,p}(x)$ converges weakly in $\mathfrak{P}_\alpha$ on G to $z_i(x)$ and $w_{i,p}(y)$ converges weakly in $\mathfrak{P}_\alpha$ on R_1 or R_2 to $w_i(y)$ $(i = 1, \cdots, N)$. From the first paragraph, it follows that $w_{i,p}(y)$ converges strongly in L_α on R_1 and R_2 to $w_i(y)$ so that $z_{i,p}(x)$ and $z_p(x)$ tend strongly in L_α on G to $z_i(x)$ and $z(x)$, respectively. The proof of (iv) of Lemma 7.3 carries over verbatim to prove that $\bar{w}_{i,p}(y)$ converges strongly in L_α on R_1^* or R_2^* to $\bar{w}_i(y)$; the step (7.2) being possible in the present case on account of the weak convergence. From this the last statement follows easily (as in the proof of Theorem 7.1).

The remainder of this section is devoted to the study of simple conditions under which families of functions in $\mathfrak{P}_\alpha$ are compact with respect to weak convergence in $\mathfrak{P}_\alpha$. If $\alpha > 1$, a necessary and sufficient condition for this is that the families be uniformly bounded in $\mathfrak{P}_\alpha$. Accordingly, Theorems 8.6 and 8.7 give conditions that a family be uniformly bounded in $\mathfrak{P}_\alpha$. The principles used in the proofs of these theorems are given in Lemma 8.4.

LEMMA 8.4. *Let $\{z(x)\}$ be a family of functions, each of class $\mathfrak{P}$ on the cell R: (a, b), with $D_\alpha(z, R)$ uniformly bounded $(\alpha \geqq 1)$; let φ denote the boundary values of z. Then all the functions z are of class $\mathfrak{P}_\alpha$ on R; and*

(a) *if $\bar{D}_\alpha(z, \tau)$ is uniformly bounded for some cell τ in R, $\bar{D}_\alpha(z, R)$ is uniformly bounded;*

(b) *if $\int_{F_k} |\varphi|^\alpha \, ds(e)$ (Lemma 7.2) is uniformly bounded for some face F_k, $x_k = a_k$ or b_k, of R^*, then $\bar{D}_\alpha(z, R)$ is uniformly bounded; and*

(c) *if $\bar{D}_\alpha(z, R)$ is uniformly bounded, $\int_{R^*} |\varphi|^\alpha \, ds(e)$ is uniformly bounded.*

To prove (a), let τ be the cell (c, d) with $a_l \leqq c_l < d_l \leqq b_l$ $(l = 1, \cdots, n)$. We define the cells $\tau_1, \cdots, \tau_n = R$ as follows: τ_1 is the cell $a_1 \leqq x_1 \leqq b_1$, $c_l \leqq x_l \leqq d_l$ $(l = 2, \cdots, n)$; τ_2 is the cell $a_1 \leqq x_1 \leqq b_1$, $a_2 \leqq x_2 \leqq b_2$, $c_l \leqq$

$x_l \leqq d_l$ $(l = 3, \cdots, n)$; and, in general, τ_m is the cell $a_k \leqq x_k \leqq b_k$ $(k = 1, \cdots, m)$, $c_l \leqq x_l \leqq d_l$ $(l = m + 1, \cdots, n)$.

Now, if $a_1 \leqq x_{1,1} \leqq c_1$ or $d_1 \leqq x_{1,1} \leqq b_1$ and $c_1 \leqq x_{1,2} \leqq d_1$, we see that

$$\int_{c_1'}^{d_1'} | \bar{z}(x_1', x_{1,2}) - \bar{z}(x_1', x_{1,1}) |^\alpha dx_1' \leqq | x_{1,2} - x_{1,1} |^{\alpha - 1} \int_R \left| \frac{\partial \bar{z}}{\partial x_1} \right|^\alpha dx$$

which is uniformly bounded for all $x_{1,1}$ in the above range and all z in the family. Thus $\bar{D}_\alpha(z, \tau_1)$ is uniformly bounded. In a similar way, one sees in succession that $\bar{D}_\alpha(z, \tau_2), \cdots, \bar{D}_\alpha(z, \tau_n) = \bar{D}_\alpha(z, R)$ are uniformly bounded.

The proofs of (b) and (c) follow at once from

$$\int_{a_k'}^{b_k'} | \bar{z}(x_k', x_{k,1}) - \bar{z}(x_k', x_{k,2}) |^\alpha dx_k' \leqq | x_{k,2} - x_{k,1} |^{\alpha - 1} D_\alpha(z, R)$$

$$(a_k \leqq x_{k,1}, x_{k,2} \leqq b_k; k = 1, \cdots, n),$$

where $\bar{z}(x_k', b_k) = \varphi(x_k', b_k)$ and $\bar{z}(x_k', a_k) = \varphi(x_k', a_k)$ and $\bar{z}$ is of class $\mathfrak{P}_\alpha'$ and equivalent to z.

THEOREM 8.6. *Let $\{z(x)\}$ be a family of functions of class $\mathfrak{P}$ on G with $D_\alpha(z, G)$ uniformly bounded. Suppose that there exists a cell R interior to G on which $\bar{D}_\alpha(z, R)$ is uniformly bounded. Then each z is of class $\mathfrak{P}_\alpha$ on any region H with $\bar{H} \subset G$ and $\bar{D}_\alpha(z, H)$ is uniformly bounded. If G is of class* K, *each z is of class $\mathfrak{P}_\alpha$ on G and $\bar{D}_\alpha(z, G)$ and $\int_{G^*} | \varphi |^\alpha ds(e)$ are uniformly bounded, φ being the boundary function for z.*

To prove the first statement, it is sufficient to prove it for cells H. Let $P \epsilon R$ and $Q \epsilon H$. It is clear that we may find a finite sequence, $R_1, \cdots, R_N$, of open cells such that R_1 contains P, R_N contains Q, and $R_i \cdot R_{i-1} \neq 0$ $(i = 1, \cdots, N + 1;$ $R_0 = R$, $R_{N+1} = H)$. That $\bar{D}_\alpha(z, H)$ is uniformly bounded follows from repeated application of (a) of Lemma 8.4.

Now, suppose G is of class K, let $\gamma_1, \cdots, \gamma_N$ be a covering of $\bar{G}$ as in Definition 7.1, let $T_1, \cdots, T_N$ be the corresponding transformations of γ_i onto R_1 or R_2, and let $w_i(y)$ be the transform under T_i of $z(x)$ as defined on γ_i. Since those γ_i with $\gamma_i \cdot G^* \neq 0$ cover all the points on and near G^*, there exists a region H with $\bar{H} \subset G$ such that $\gamma_i \cdot H$ is open for each i. Since $\bar{D}_\alpha(z, H)$ is uniformly bounded, and since the transform of $\gamma_i \cdot H$ under T_i contains a cell, it follows from Lemma 8.4 (a) that $\bar{D}_\alpha(R_1, w_i)$ and $\bar{D}_\alpha(R_2, w_i)$ are uniformly bounded so that, by addition, $\bar{D}_\alpha(z, G)$ is uniformly bounded. That $\int_{G^*} | \varphi |^\alpha ds(e)$ is uniformly bounded follows from (c) of Lemma 8.4.

THEOREM 8.7. *Let $\{z(x)\}$ be a family of functions of class $\mathfrak{P}_\alpha$ $(\alpha \geqq 1)$ on a region G, of class* K, *with $D_\alpha(z, G)$ uniformly bounded. Suppose that there exists*

a set E, open on G^*, such that $\int_E |\varphi|^\alpha ds(e)$ *is uniformly bounded. Then* $\bar{D}_\alpha(z, G)$ *is uniformly bounded.*

From Definition 7.1 it follows that there exists a subset γ of $\bar{G}$, with $\gamma \cdot G^* \neq 0$ and $\gamma \cdot G^* \subset E$, which is the image, under the regular transformation $T: x = x(y)$, of the cell R_2, the points of $\gamma \cdot G^*$ and those of R_2 where $y_n = 0$ corresponding under T. Let $w(y)$ be the transform of $z(x)$ and $\psi(y_n') = \varphi[x(y_n', 0)]$. Then $\bar{w}(y_n', 0) = \psi(y_n')$ almost everywhere and $\int_{-1}^{1} |\psi(y_n')|^\alpha dy_n'$ is uniformly bounded. From (b) of Lemma 8.4, it follows that $\bar{D}_\alpha(w, R_2)$ and hence $\bar{D}_\alpha(z, \gamma)$ is uniformly bounded. Since γ contains a cell H with $\bar{H} \subset G$, the theorem follows from Theorem 8.6.

If $\alpha = 1$, the mere uniform boundedness of $\bar{D}_\alpha(z, G)$ is not sufficient for $\{z\}$ to be compact with respect to weak convergence. The following theorem gives an interesting condition for compactness in this case. (Cf. [7].)

THEOREM 8.8. *A necessary and sufficient condition that the family* $\{z(x)\}$ *of functions of class* $\mathfrak{P}_1$ *on the bounded region* G *be compact with respect to weak convergence in* $\mathfrak{P}_1$ *on* G *is that the following two conditions hold:*

(i) $\bar{D}_1(z, G)$ *is uniformly bounded;*

(ii) *there exists a non-negative convex*[4] *function* $\varphi(r_1, \cdots, r_n)$ *with the property that*

$$\lim_{|r| \to \infty} |r|^{-1} \varphi(r_1, \cdots, r_n) = +\infty, \qquad |r|^2 = r_1^2 + \cdots + r_n^2, \tag{8.6}$$

and such that

$$\int_G \varphi[D_{x_1} z, \cdots, D_{x_n} z]\, dx \tag{8.7}$$

is uniformly bounded.

From Theorems 8.2 and 8.4, it follows that a necessary and sufficient condition for this compactness is that (i) should hold and that the set functions

$$\int_e \left[\sum_{i=1}^{n} D_{x_i}^2 z\right]^{\frac{1}{2}} dx \equiv D_1(z, e)$$

should be uniformly absolutely continuous.

Hence, let us assume that these conditions are satisfied in our family and let M be a uniform bound for $\bar{D}_1(z, G)$. For each p and z, let $E_p(z)$ be the set of x where $p - 1 < \left[\sum_{i=1}^{n} D_{x_i}^2 z\right]^{\frac{1}{2}} \leq p$ and let $\mathcal{E}_p(z) = \sum_{k=p+1}^{\infty} E_k(z) + Z(z)$, $Z(z)$ being the set of measure zero where one of the $D_{x_i} z$ fails to exist; then $\mathcal{E}_p(z)$

[4] A function $f(x)$ is convex if $f[(1 - \lambda)x_1 + \lambda x_2] = (1 - \lambda)f(x_1) + \lambda f(x_2)$ for each λ, $0 \leq \lambda \leq 1$, and each pair of points (x_1, x_2).

is the set where $\left[\sum_{i=1}^{n} D_{x_i}^2 z\right]^{\frac{1}{2}} > p$ or is not defined. Clearly $m[\mathcal{E}_p(z)] < M/p$ so that

$$D_1[z, \mathcal{E}_p(z)] \leqq \sum_{k=p+1}^{\infty} k \cdot m[E_k(z)] \leqq D_1[z, \mathcal{E}_p(z)] + m[\mathcal{E}_p(z)] \leqq \epsilon_p, \qquad \lim_{p\to\infty} \epsilon_p = 0,$$

ϵ_p being independent of z on account of the uniform absolute continuity. Now let p_1 be the first integer such that $\epsilon_p \leqq \frac{1}{2}$ for $p \geqq p_1$, let p_2 be the first integer $> p_1$ such that $\epsilon_p \leqq (\frac{1}{2})^2$ for all $p \geqq p_2$, and, in general, let p_{k+1} be the first integer $> p_k$ such that $\epsilon_p \leqq 2^{-k-1}$ for all $p \geqq p_{k+1}$. Define $\psi(|r|)$ by

$$\psi(|r|) = \begin{cases} |r|, & 0 \leqq |r| \leqq p_1, \\ 2|r| - p_1, & p_1 \leqq |r| \leqq p_2, \\ 3|r| - p_1 - p_2, & p_2 \leqq |r| \leqq p_3, \\ \cdots\cdots\cdots\cdots\cdots\cdots & \\ k|r| - \sum_{i=1}^{k-1} p_i, & p_{k-1} \leqq |r| \leqq p_k, \\ \cdots\cdots\cdots\cdots\cdots\cdots . & \end{cases}$$

Then $\psi(|r|)$ is convex in $|r|$ and we have

$$\psi(|r|) \geqq \left[1 + \left(1 - \frac{p_1}{p_k}\right) + \left(1 - \frac{p_2}{p_k}\right) + \cdots + \left(1 - \frac{p_{k-1}}{p_k}\right)\right] \cdot |r|, \quad |r| \geqq p_{k+1}.$$

We define $\varphi(r_1, \cdots, r_n) = \psi[(r_1^2 + \cdots + r_n^2)^{\frac{1}{2}}] = \psi(|r|)$ and φ is convex and satisfies (8.6). Also

$$\int_G \varphi[D_{x_1} z, \cdots, D_{x_n} z]\, dx \leqq D_1(z, G) + \sum_{k=1}^{\infty} k \cdot D_1[z, \mathcal{E}_{p_k}(z)] \leqq M + \sum_{k=1}^{\infty} k \cdot 2^{-k}$$

independently of z.

Next, we assume that the conditions of the theorem are satisfied and take M as a uniform bound for the integral (8.7). For each p, let α_p be the greatest lower bound of $(r_1^2 + \cdots + r_n^2)^{-\frac{1}{2}} \cdot \varphi(r_1, \cdots, r_n)$ for $(r_1^2 + \cdots + r_n^2)^{\frac{1}{2}} \geqq p - 1$; then $\lim_{p\to\infty} \alpha_p = +\infty$. Then, for each p,

$$\sum_{k=p+1}^{\infty} \alpha_k \cdot (k-1) m[E_k(z)] \leqq \int_G \varphi[D_{x_1} z, \cdots, D_{x_n} z]\, dx \leqq M$$

for each z in $\{z\}$. Hence

$$D_1[z, \mathcal{E}_p(z)] \leqq \sum_{k=p+1}^{\infty} k \cdot m[E_k(z)] \leqq M(p+1)/p\alpha_{p+1}, \qquad \lim_{p\to\infty} M(p+1)/p\alpha_p = 0,$$

the bound being independent of z. Now choose $\epsilon > 0$. Let p be so large that $M(p+1)/p\alpha_p < \frac{1}{2}\epsilon$. For that p, choose $\delta = \epsilon/2p$. Then if $m(e) < \delta$, we see that $D_1(z, e) < \epsilon$ for each z in the family.

It is an obvious but convenient remark that Theorem 8.8 holds if $\alpha > 1$; we need only to take $\varphi = (r_1^2 + \cdots + r_n^2)^{\frac{1}{2}\alpha}$.

9. **Boundary values on general regions.** In this section, we discuss functions of class $\mathfrak{P}_\alpha$, defined on an arbitrary *bounded* region G, particularly with respect to their boundary values. It is clear that we will ordinarily have no boundary value function φ as in §7. However, (iv) of Theorem 7.1 suggests the definition given below for the vanishing of a function of class $\mathfrak{P}_\alpha$ on G^*. This definition has been given for $\alpha = 2$ in volume 2, Chapter VII, §1 of Courant-Hilbert, *Methoden der mathematischen Physik*.

DEFINITION 9.1. We say that a function z of class $\mathfrak{P}_\alpha$ *vanishes* (or $= 0$) *on* G^* if there exists a sequence $\{z_p(x)\}$ which converges strongly in $\mathfrak{P}_\alpha$ on G to z, each z_p satisfying a uniform Lipschitz condition on $\bar{G}$ and vanishing for x not in a subregion G_p of G with $\bar{G}_p \subset G$, where any closed (bounded) subset F of G is in all of the G_p for $p > P_F$.

If G is of class K, this definition is equivalent (in view of Theorem 7.1) to the condition that the boundary function vanish. We first prove a theorem analogous to Theorem 7.3 for the case that $z(x)$ vanishes on G^*. It can be proved that the necessary condition mentioned in the theorem is sufficient as well as necessary. This proof requires a great deal more analysis than the degree of interest of the result warrants. For simplicity of statement, we interpose the following definition.

DEFINITION 9.2. A function $z(x)$ is *of class* $\mathfrak{P}'''$ on a region G if it is of class $\mathfrak{P}'$ there and if $z[x(y)]$ is of class $\mathfrak{P}'$ in y on H if $x = x(y)$ is any regular transformation of class K of H into G (H being variable with the transformation); $z(x)$ is *of class* $\mathfrak{P}_\alpha'''$ if it is also of class $\mathfrak{P}_\alpha$.

If $z(x)$ is of class $\mathfrak{P}_\alpha$ on G, then the function $\bar{z}(x)$ of (6.8) is of class $\mathfrak{P}_\alpha'''$ on G.

THEOREM 9.1. *Suppose that $z(x)$ is of class $\mathfrak{P}_\alpha$ on G. Then a necessary condition that $z(x)$ vanish on G^* is that the function $Z(x) = \bar{z}(x)$ on G and zero elsewhere be of class $\mathfrak{P}_\alpha'''$ on any bounded region of space. If $z(x)$ vanishes on G^* and $x = x(y)$ is a regular transformation, of class* K, *of a region H into G and if $w(y) = z[x(y)]$, then $w(y)$ vanishes on H^*.*

The second statement is an immediate consequence of §6 and Definition 9.1. To prove the first statement, let H be any region of class K such that $H \supset \bar{G}$. Let $\{z_p(x)\}$ be a sequence as in Definition 9.1 and extend each $z_p(x)$ and $z(x)$ to H by defining each to be zero on $H - G$. Then $z(x)$ is of class $\mathfrak{P}_\alpha$ on H and $z_p(x)$ is a sequence for z on H as in Definition 9.1. As in the proof of Theorem 7.3, we show that, for each k ($1 \leqq k \leqq n$), there exists a subset Z_k of $(n - 1)$-dimensional measure zero such that if $\bar{x}_k'$ is not in Z_k, then a subsequence $\{z_q(\bar{x}_k', x_k)\}$ converges uniformly in x_k on each closed segment of $x_k' = \bar{x}_k'$ the interior of which is in H to $\bar{z}(\bar{x}_k', x_k)$, the function of (6.8) for $z(x)$ on H. Since $z_p(x) = 0$ for each p and each x in $H - G$, it follows that the above function $Z(x)$ is of class $\mathfrak{P}_\alpha'$ on H. It is easily seen, by making a regular transforma-

tion $x' = x'(x)$ of H into H', that $Z(x)$ is of class $\mathfrak{P}_\alpha'''$ on H. This proves the theorem.

Definition 9.1 suggests the following definition.

DEFINITION 9.3. Let z_1 and z_2 be of class $\mathfrak{P}_\alpha$ on G. We say that z_1 *and* z_2 *coincide* (or $z_1 = z_2$) *on* G^* if $z_1 - z_2$ vanishes on G^* in the sense of Definition 9.1.

We now prove a fundamental theorem involving weak convergence in $\mathfrak{P}_\alpha$ and the above concepts of boundary values.

THEOREM 9.2. *Let* $\{z_p(x)\}$ *converge weakly in* $\mathfrak{P}_\alpha$ *on* G *to* $z(x)$ *and suppose that* $z_p(x) = z^*(x)$ *on* G^* $(p = 1, 2, \cdots)$, $z^*(x)$ *being of class* $\mathfrak{P}_\alpha$ *on* G. *Then* $z(x) = z^*(x)$ *on* G^*.

Clearly there is no loss of generality in assuming that $z^*(x) \equiv 0$. For each $z_p(x)$, let $Z_p(x)$ satisfy a uniform Lipschitz condition on $\bar{G}$, vanish on and near G^*, and satisfy

$$\bar{D}(z_p - Z_p, G) < \frac{1}{p}.$$

Then it is clear (since strong convergence implies weak convergence) that $\{Z_p(x)\}$ tends weakly in $\mathfrak{P}_\alpha$ to $z(x)$ on G. From a well-known theorem ([1], Chapter IX) on weak convergence, it follows that there exists a sequence $\{(c_{p,1}Z_1 + \cdots + c_{p,p}Z_p)\}$ which converges strongly in $\mathfrak{P}_\alpha$ to $z(x)$ on G. This sequence is evidently a sequence for $z(x)$ as in Definition 9.1, so that $z(x) = 0$ on G^*.

We now develop a convenient representation and an important inequality for functions of class $\mathfrak{P}_\alpha$ which vanish on G^*, G being bounded. We state these results in Theorem 9.3 below. In the following, we define γ_n and Γ_n by the formulas

$$m[C(P, r)] = \gamma_n r^n, \qquad m[C^*(P, r)] = \Gamma_n r^{n-1},$$

the second being the $(n-1)$-dimensional hyperarea. It is clear that $\Gamma_n = n \cdot \gamma_n$. We first prove a simple lemma which, however, is important for the sequel.

LEMMA 9.1. *If* E *is any set of finite measure and* x_0 *is any point, then*

$$\int_E \left[\sum_{i=1}^n (x_i - x_{i,0})^2\right]^{-\frac12(n-1)} dx \leqq \int_{C(x_0,r)} \left[\sum_{i=1}^n (x_i - x_{i,0})^2\right]^{-\frac12(n-1)} dx = \Gamma_n r$$
$$= \Gamma_n[m(E)/\gamma_n]^{1/n},$$
$$\gamma_n r^n = m(E).$$

The first inequality is immediate, since the integrand is smaller for any x not in $C(x_0, r)$ than for any x in $C(x_0, r)$ so the integral is increased by substituting $C(P_0, r) - C(P_0, r) \cdot E$ for the part, which is of equal measure, of E not in $C(P_0, r)$. The remaining steps follow by integration and substitution.

THEOREM 9.3. *Let* z *be of class* $\mathfrak{P}_\alpha$ *on the bounded region* G *and vanish on* G^*. *If we define* $z = 0$ *for* x *not in* G, *then*

$$\text{(9.1)} \qquad \bar{z}(x) = -\Gamma_n^{-1} \int_G \left[\sum_{i=1}^n (\xi_i - x_i)^2\right]^{-\frac12 n} \cdot \left[\sum_{i=1}^n (\xi_i - x_i) p_i(\xi)\right] d\xi, \qquad p_i = D_{x_i} z,$$

for almost all x in space, $\bar{z}(x)$ denoting the function (6.8) *where the sets e are restricted to be spheres $C(x, r)$. In fact $\bar{z}(x)$ is defined for each point x for which the integral on the right exists as a Lebesgue integral, this being almost everywhere. Also*

$$\int_E |z(x)|^\alpha dx \leqq \gamma_n^{-\alpha/n}\cdot[m(G)]^{(\alpha-1)/n}\cdot[m(E)]^{1/n}\cdot D_\alpha(z, G) \tag{9.2}$$

for any measurable set E of finite measure.

We first observe that the integrand on the right in (9.1) is measurable in the $2n$-dimensional (x, ξ)-space as are the integrands in (9.3) below. Also the iterated integrals

$$\begin{aligned}(9.3)\quad &\int_G \left[\sum_{i=1}^n p_i^2(\xi)\right]^{\frac{1}{2}\beta}\left\{\int_E\left[\sum_{i=1}^n(\xi_i - x_i)^2\right]^{-\frac{1}{2}(n-1)}dx\right\}d\xi \\ &\leqq \Gamma_n\cdot[m(E)/\gamma_n]^{1/n}\cdot D_\beta(z, G) \qquad (1 \leqq \beta \leqq \alpha;\ p_i = D_{x_i}z)\end{aligned}$$

all exist if E is any measurable set of finite measure. Thus the integral in (9.1) exists as a Lebesgue integral for almost every x.

Now, let x be a point where this integral converges and let (r, σ) be polar coördinates with pole at x and let $\bar{w}(r, \sigma)$ be the transform of $\bar{z}(\xi)$ (see (6.8)), σ denoting a point on the unit sphere $C^*(0, 1)$. Then $\bar{w}$ is of class $\mathfrak{P}_\alpha'''$ in any bounded region of (r, σ)-space in which $r \geqq r_0 > 0$ and the usual differentiation formulas hold and

$$\begin{aligned}(9.4)\quad &\int_{C(x,R)-C(x,\rho)}\left[\sum_{i=1}^n(\xi_i - x_i)^2\right]^{-\frac{1}{2}n}\left[\sum_{i=1}^n(\xi_i - x_i)p_i(\xi)\right]d\xi \\ &= \int_\rho^R\int_{C^*(0,1)}\frac{\partial\bar{w}}{\partial r}dr\,d\sigma = \int_{C^*(0,1)}[\bar{w}(R, \sigma) - \bar{w}(\rho, \sigma)]\,d\sigma \qquad (p_i = D_{x_i}z).\end{aligned}$$

Since the above integral is absolutely convergent as $\rho \to 0$, we see that there exists a summable function $\bar{w}(0, \sigma)$ such that

$$\lim_{r\to 0}\int_{C^*(0,1)}|\bar{w}(r, \sigma) - \bar{w}(0, \sigma)|\,d\sigma = 0.$$

From this, it follows easily, by choosing R so large that $C(x, R)$ contains $\bar{G}$, that

$$\bar{z}(x) = \Gamma_n^{-1}\int_{C^*(0,1)}\bar{w}(0, \sigma)\,d\sigma = -\Gamma_n^{-1}\int_0^R\int_{C^*(0,1)}\frac{\partial\bar{w}}{\partial r}dr\,d\sigma$$

from which the existence of $\bar{z}(x)$ and formula (9.1) follows.

Now, from the above it follows that, for almost all x,

$$\begin{aligned}|z(x)| \leqq \Gamma_n^{-1}\int_G &\left\{\left[\sum_{i=1}^n(\xi_i - x_i)^2\right]^{-\frac{1}{2}(n-1)/\alpha}\right. \\ &\left.\cdot\left[\sum_{i=1}^n p_i^2(\xi)\right]^{\frac{1}{2}}\right\}\cdot\left\{\left[\sum_{i=1}^n(\xi_i - x_i)^2\right]^{-\frac{1}{2}(n-1)(\alpha-1)/\alpha}\right\}d\xi\end{aligned}$$

$$\leqq \Gamma_n^{-1}\left\{\int_G \left[\sum_{i=1}^n (\xi_i - x_i)^2\right]^{-\frac{1}{2}(n-1)} \cdot \left[\sum_{i=1}^n p_i^2(\xi)\right]^{\frac{1}{2}\alpha} d\xi\right\}^{1/\alpha}\left\{\int_G \left[\sum_{i=1}^n (\xi_i - x_i)^2\right]^{-\frac{1}{2}(n-1)} d\xi\right\}^{(\alpha-1)/\alpha}$$

$$\leqq \Gamma_n^{-1}\{\Gamma_n\cdot[m(G)/\gamma_n]^{1/n}\}^{(\alpha-1)/\alpha}\left\{\int_G\left[\sum_{i=1}^n (\xi_i - x_i)^2\right]^{-\frac{1}{2}(n-1)}\left[\sum_{i=1}^n p_i^2(\xi)\right]^{\frac{1}{2}\alpha} d\xi\right\}^{1/\alpha}$$

$$(p_i = D_{x_i} z).$$

(9.2) then follows by integrating with respect to x, changing the order of integration, and then using (9.3) with $\beta = \alpha$.

We conclude this section with a result for general regions, corresponding roughly to Theorem 8.7. This theorem, together with Theorem 8.8, gives a condition that a family of functions in $\mathfrak{P}_\alpha$ on G be compact with respect to weak convergence. This theorem is an immediate consequence of Theorems 9.2 and 9.3.

THEOREM 9.4. *Let $\{z^*\}$ be any family of functions uniformly bounded in $\mathfrak{P}_\alpha$ on G. Let $\{z\}$ be any family of functions of class $\mathfrak{P}_\alpha$ on G with $D_\alpha(z, G)$ uniformly bounded, each z coinciding on G^* with some z^* in $\{z^*\}$.*

Then $\bar{D}_\alpha(z, G)$ is uniformly bounded.

BIBLIOGRAPHY

1. S. BANACH, *Théorie des Opérations Linéaires*, Warsaw, 1932.
2. J. W. CALKIN, *Functions of several variables and absolute continuity*, I, this Journal, vol. 6(1940), pp. 170–186.
3. G. C. EVANS, *Complements of Potential Theory* (II), American Journal of Mathematics, vol. 55(1933), pp. 29–49.
4. G. C. EVANS, *Fundamental Points of Potential Theory*, Rice Institute Pamphlets, vol. VII(1920), no. 4, pp. 252–329.
5. G. C. EVANS, *Note on a theorem of Bôcher*, American Journal of Mathematics, vol. 50(1928), pp. 123–126.
6. K. FRIEDRICHS, *Die Randwert- und Eigenwertprobleme aus der Theorie der elastischen Platten*, Mathematische Annalen, vol. 98(1927), pp. 205–247.
7. M. NAGUMO, *Über die gleichmässige Summierbarkeit und ihre Anwendung auf ein Variationsproblem*, Japanese Journal of Mathematics, vol. 6(1929), pp. 173–182.
8. H. RADEMACHER, *Über partielle und totale Differenzierbarkeit von Funktionen mehrerer Variabeln, und über die Transformation der Doppelintegrale*, Mathematische Annalen, vol. 79(1918), pp. 340–359.
9. F. RELLICH, *Ein Satz über mittlere Konvergenz*, Göttingen Nachrichten (Math.-Phys. Klasse), (1930), pp. 30–35.
10. L. TONELLI, *Sulla quadratura delle superficie*, Atti della Reale Accademia dei Lincei, (6), vol. 3(1926), pp. 633–638.
11. C. J. DE LA VALLÉE POUSSIN, *Intégrales de Lebesgue, Fonctions d'Ensemble, Classes de Baire*, Paris, 1916.
12. O. VEBLEN AND J. H. C. WHITEHEAD, *The Foundations of Differential Geometry*, Cambridge, 1932.

UNIVERSITY OF CALIFORNIA.

EXTENDING MAPS OF PLANE PEANO CONTINUA

BY V. W. ADKISSON AND SAUNDERS MAC LANE

1. **Introduction.** When can a homeomorphism T of a Peano continuum M on a sphere S to a set M' on a sphere S' be extended to a homeomorphism T' which carries the whole sphere S into S'? Gehman has solved this problem for the extension of T to a map of a *plane* containing such a space M.[1] His condition can easily be modified to include the case of M and M' on spheres.[2] But this condition, which requires that certain "sides" of M be preserved by T, is extrinsic to the given M, and requires in fact that one establish the existence of certain arcs of S and S' which may cut through M and its complement $S - M$ in a complicated fashion. We establish here another necessary and sufficient condition for the extendibility of T. This condition is intrinsic, and applies only to the triods of M, where by a *triod* in M we mean a set $\tau = [\alpha, \beta, \gamma]$ of three arcs of M with a common end point, the vertex of the triod, such that any two of these arcs intersect only in this vertex.

THEOREM 1. *If M and M' are topologically equivalent Peano continua lying respectively on spheres S and S', then a homeomorphism T of M to M' can be extended to a homeomorphism T' of S to S' if and only if T preserves the relative sense of every pair of triods of M.*

To say that T *preserves the relative sense of triods* here means that T carries any two triods τ_1 and τ_2 of M which have the same sense on S into two triods τ_1' and τ_2' which have the same sense (i.e., both are clockwise or both are counterclockwise) on S'. The precise method for treating this concept of "sense" is sketched in §2.

This theorem implies that when T preserves sense on triods, it necessarily carries each complementary domain boundary (c.d.b.) of M into a c.d.b. of M'. An extendibility condition for planes may thus be found by projecting the sphere from a point in a suitable complementary domain.

THEOREM 2. *A homeomorphism $T(M) = M'$ between two plane Peano continua M and M' can be extended to the whole planes if and only if T preserves the*

Received September 25, 1939; presented to the American Mathematical Society, October 28, 1939.

[1] H. M. Gehman, *On extending a continuous* (1-1) *correspondence of two plane continuous curves to a correspondence of their planes,* Transactions of the American Mathematical Society, vol. 28(1926), pp. 252–265.

[2] V. W. Adkisson, *On extending a continuous* (1-1) *correspondence of continuous curves on a sphere,* Comptes Rendus des Séances de la Société des Sciences et des Lettres de Varsovie, vol. 27(1934), pp. 5–9.

H. M. Gehman, *On extending a homeomorphism between two subsets of spheres,* Bulletin of the American Mathematical Society, vol. 42(1936), pp. 79–81.

relative sense of every pair of triods of M, and also if for the boundary k of the unbounded complementary domain of M the complementary domain boundary $T(k)$ of M' bounds the unbounded complementary domain of M'.

The structure of the set M may be used to obtain more refined conditions for extendibility. For a cyclically connected M one may easily prove, by the Schoenflies Theorem, that $T(M) = M'$ can be extended to the spheres if and only if T carries each c.d.b. of M into a c.d.b. of M'. If M is triply connected, then any homeomorphism T whatever can be extended to the spheres.[3]

Definition. A continuum M is *triply connected*[4] if M contains no split pair of points p and q.

Definition. Two points p and q of a continuum M constitute a *split pair* of M if $M - (p + q)$ is the sum of two mutually separated sets H_1 and H_2, neither of which is an arc from p to q.

The behavior of T inside triply connected subsets of M does not interfere with the extendibility of T as is stated in

THEOREM 3. *The sufficient condition for the extendibility of $T(M) = M'$, as given in Theorems* 1 *and* 2, *remains valid if it is required not for all triods in M, but only for the triods of any list which includes*

(a) *at each cut point P of M, all triods with vertex P,*

(b) *for each split pair p, q of a proper cyclic element of M, all triods with vertex at one of these two points (say at p),*

(c) *for each triply connected proper cyclic element N of M, any one triod in N.*

The proof of this theorem involves a process of imbedding a θ-graph contained in M in a *maximal triply connected subcontinuum* of M. The structure of M relative to such maximal elements suggests further generalizations of Whyburn's cyclic element theory.[5] The efficiency of our triod condition for further investigations of extendibility may be observed by noting the simplicity with which it could have been applied to investigations of continua with unique maps.[6]

2. Sense on triods. The sense of a triod $\tau = [\alpha, \beta, \gamma]$ on a sphere S can be defined because S is an orientable manifold. Any topological map $T(S)$ of the sphere on itself either preserves the orientation, and then has the degree (Abbildungsgrad) $+1$, or inverts the orientation of the sphere, and then has degree -1. The maps of degree $+1$ can also be characterized as the maps which can be homotopically deformed into the identity.

[3] V. W. Adkisson, *Plane peanian continua with unique maps on the sphere and in the plane*, Transactions of the American Mathematical Society, vol. 44(1938), p. 59, Theorem II.

[4] Triply connected graphs are discussed by Whitney, *Congruent graphs and the connectivity of graphs*, American Journal of Mathematics, vol. 54(1932), p. 158. See also Mac Lane and Adkisson, *Fixed points and the extension of the homeomorphisms of a planar graph*, American Journal of Mathematics, vol. 60(1938), pp. 611–639.

[5] G. T. Whyburn, *Concerning the structure of a continuous curve*, American Journal of Mathematics, vol. 50(1928), pp. 167–194.

[6] V. W. Adkisson, *Plane peanian continua*, loc. cit.

Definition. Two triods $\tau_1 = [\alpha_1, \beta_1, \gamma_1]$ and $\tau_2 = [\alpha_2, \beta_2, \gamma_2]$ on a sphere S have the *same sense* on S if and only if there is a topological map $T(S)$ of degree 1 of S into itself, which takes τ_1 into τ_2, α_1 into α_2, β_1 into β_2, and γ_1 into γ_2. Two triods which do not have the same sense are said to have *opposite sense.*

We need to compare this definition with Kline's definition of sense on simple closed curves.[7] In proving properties of sense it is convenient to observe that S can be topologically mapped onto another sphere S' so that a given triod τ goes into a triod τ' in some standard position,[8] such as the position with vertex at the north pole and legs α', β', γ' arcs of great circles. Using this standard position, one observes at once that the triods $[\alpha, \beta, \gamma]$, $[\beta, \gamma, \alpha]$ and $[\gamma, \alpha, \beta]$ have the same sense, and that "equivalent" triods have the same sense.

Definition. Two triods $\tau_1 = [\alpha_1, \beta_1, \gamma_1]$ and $\tau_2 = [\alpha_2, \beta_2, \gamma_2]$ with the same vertex are *equivalent* if there exists a triod $\tau_3 = [\alpha_3, \beta_3, \gamma_3]$ with the same vertex, such that $\alpha_3 \subset \alpha_1 \cdot \alpha_2$, $\beta_3 \subset \beta_1 \cdot \beta_2$, $\gamma_3 \subset \gamma_1 \cdot \gamma_2$.

THEOREM 4. *The triods $[\alpha, \beta, \gamma]$ and $[\alpha, \gamma, \beta]$ have opposite sense.*

One need only prove that any map T taking the first triod into the second must have degree -1. To show this, add enough arcs to the triod $[\alpha, \beta, \gamma]$ to make it part of a simplicial subdivision K of the sphere. The map T takes K into another simplicial subdivision K'. But T interchanges β and γ and leaves α fixed, hence T maps K on K' with inversion of orientation.

By suitable deformations of S one may prove

THEOREM 5. *The triods $[\alpha_1, \beta_1, \gamma_1]$ and $[\alpha_2, \beta_2, \gamma_2]$ have opposite senses if and only if $[\alpha_1, \beta_1, \gamma_1]$ and $[\alpha_2, \gamma_2, \beta_2]$ have the same sense.*

From this it follows that when τ_1 and τ_2 have opposite sense, while τ_2 and τ_3 also have opposite sense, then τ_1 and τ_3 have the same sense.

THEOREM 6. *Two non-intersecting triods $\tau_1 = [\alpha_1, \beta_1, \gamma_1]$ and $\tau_2 = [\alpha_2, \beta_2, \gamma_2]$ have opposite sense on a sphere S if and only if there exists on S a θ-graph whose vertices are the vertices of τ_1 and τ_2 and whose three (non-intersecting) arcs contain respectively the legs α_1 and α_2, β_1 and β_2, γ_1 and γ_2.*

Proof. If there exist three such independent arcs constituting a θ-graph, this graph may be topologically distorted into standard position (say arcs of great circles). In this standard position one may readily construct a map of degree $+1$ which takes α_1, β_1, γ_1 into α_2, γ_2, β_2 respectively. Hence τ_1 and τ_2 have opposite sense by Theorem 5. Conversely, given τ_1 and τ_2, one may

[7] J. R. Kline, *A definition of sense on closed curves in non-metrical plane analysis situs,* Annals of Mathematics, vol. 19(1918), pp. 185–200. See also J. R. Kline, *Concerning sense on closed curves in non-metrical plane analysis situs,* Annals of Mathematics, vol. 21(1919), pp. 113–119.

[8] The existence of the map of S on S' which carries τ in specified fashion into τ' is a special instance of an extension of a homeomorphism. In each case the existence of such an extension can readily be proved by drawing additional arcs on S and S' (so as to make τ and τ' parts of a simplicial subdivision), then applying the Schoenflies Theorem.

construct a simple closed curve J on S containing α_1, α_2, β_1 and β_2, but not otherwise intersecting γ_1 and γ_2. If γ_1 and γ_2 lie on opposite sides of J, one may distort J into standard position and then construct a map of degree $+1$ which takes τ_1 into τ_2. This makes τ_1 and τ_2 have the same sense, contrary to hypothesis. Therefore γ_1 and γ_2 must lie on the same side of J, so that they can be joined by an arc which, together with J, constitutes the desired θ-graph. This argument also proves

THEOREM 7. *Let $\tau_1 = [\alpha_1, \beta_1, \gamma_1]$ and $\tau_2 = [\alpha_2, \beta_2, \gamma_2]$ be two non-intersecting triods on S, and let J be a simple closed curve of S containing α_1, α_2, β_2, β_1, in this order, but not meeting γ_1 and γ_2 (except at the vertices of τ_1 and τ_2). Then τ_1 and τ_2 have the same sense if and only if γ_1 and γ_2 lie on opposite sides of J.*

These two criteria will be used repeatedly. They state the essence of Kline's original definition of sense on closed curves in the plane.[9] His definition implies that two triples of points A_i, B_i, C_i on the non-intersecting (simple) closed curves J_i ($i = 1, 2$) have opposite sense (with respect to the common exterior of J_1 and J_2) if and only if this exterior contains three non-intersecting arcs A_1A_2, B_1B_2 and C_1C_2. These three arcs, added to suitable triods within J_1 and J_2, give a θ-graph. This proves

THEOREM 8. *Let $\tau_1 = [\alpha_1, \beta_1, \gamma_1]$ and $\tau_2 = [\alpha_2, \beta_2, \gamma_2]$ be two disjoint triods, while J_1 and J_2 are disjoint closed curves such that J_i separates the vertex of τ_i from that of τ_{i+1} ($i + 1$ reduced* mod 2), *and such that J_i meets α_i, β_i and γ_i. Let A_i, B_i and C_i denote respectively the first points of α_i, β_i, γ_i on J_i. Then τ_1 and τ_2 have the same sense if and only if the triple $A_1B_1C_1$ on J_1 has the same sense (with respect to the common exterior of J_1 and J_2) as the triple $A_2B_2C_2$ on J_2.*

Such a reduction of sense on triods to sense of triples of points on closed curves is valid whether the sense of these triples is defined by Kline's analysis situs method or by combinatorial methods.

As an analogue to Theorems 6 and 7 for intersecting triods we have

THEOREM 9. *Let $\tau_1 = [\alpha_1, \beta_1, \gamma_1]$ and $\tau_2 = [\alpha_2, \beta_2, \gamma_2]$ be two triods with the same vertex V, and let J be an oriented simple closed curve with interior a neighborhood of V, such that J intersects all six arcs of τ_1 and τ_2. Let A_i, B_i, C_i be respectively the first points of α_i, β_i and γ_i (from V) which lie on J. Then τ_1 and τ_2 have the same sense if and only if $A_1B_1C_1$ and $A_2B_2C_2$ have the same sense on J, in the given orientation of J.*

3. **First main theorem** (Theorem 1). Throughout this section let M and M' denote homeomorphic Peano continua on spheres S and S' respectively, and let T denote a homeomorphism $T(M) = M'$.

Definition. Let J be any simple closed curve of M. If for any two points P_1 and P_2 of $M - J$ that lie in one of the regions of S bounded by J the points

[9] Kline, loc. cit.

$T(P_1)$ and $T(P_2)$ lie in one of the regions of S' bounded by $T(J)$, then T is said to *preserve regions in* M.

THEOREM 10. *If T preserves relative sense on every pair of triods in M, then T preserves regions in M.*

Proof. Let J be any closed curve of M and $J' = T(J)$. (Primes will indicate the corresponding image under T.) Let P_1 and P_2 represent any two points of $M - J$ that lie in R, a complementary domain of J. Then P_1' and P_2' must lie in the same complementary domain of J'. Suppose the contrary; i.e., P_1' and P_2' are separated by J'. Then obviously P_1 and P_2 cannot belong to the same component of $M - J$, and there exists an arc $P_1Q_1 = \alpha_1$ of M lying in R except for its end point Q_1 on J and an arc $P_2Q_2 = \alpha_2$ of M lying in R except for Q_2 on J and independent of α_1. If $Q_1 = Q_2$, let β_1 and β_2 denote two subarcs of J meeting only at their common end point Q_1. Then it follows from the definition of sense on triod pairs and the fact that α_1 and α_2 are in the same complementary domain of J that sense $\beta_1\beta_2\alpha_1$ is the same as sense $\beta_1\beta_2\alpha_2$. But in the second map α_1' and α_2' are separated by J' so that sense $\beta_1'\beta_2'\alpha_1'$ is opposite to the sense $\beta_1'\beta_2'\alpha_2'$, as one may see by applying Theorem 9. If $Q_1 \neq Q_2$, each of the two arcs of J joining Q_1 to Q_2 may be cut into two subarcs $\beta_1 + \beta_2$ and $\gamma_1 + \gamma_2$ so that $\beta_i \cdot \gamma_i = Q_i$ $(i = 1, 2)$. Then from Theorem 7 we have the sense $\alpha_1\beta_1\gamma_1$ opposite to the sense $\alpha_2\beta_2\gamma_2$. In the second map α_1' and α_2' are separated by J' so that the sense $\alpha_1'\beta_1'\gamma_1'$ is the same as the sense $\alpha_2'\beta_2'\gamma_2'$ (Theorem 7). This contradiction proves the theorem.

THEOREM 11. *If k is a complementary domain boundary of M, then $T(k)$ is a complementary domain boundary of M' if T and T^{-1} preserve regions in M.*

Proof. Let k represent a c.d.b. of M. If P_1 and P_2 are two points of k, then P_1' and P_2' lie on a common c.d.b. of M'. For if not, then there must exist a closed curve of M' separating P_1' and P_2'.[10] Since regions are preserved, the corresponding closed curve in M must then separate the original points P_1 and P_2, contrary to the assumption that P_1 and P_2 lie on one k.

The demonstration that any two points of k' lie on a c.d.b. of M' does not of itself insure that all points of k' lie on the same c.d.b. of M'. To prove the latter statement, pick any closed curve J in k (if k contains no closed curve, then $M = k$). The set $k - J$ must lie in one of the complementary domains R of J. Since regions are preserved, $k' - J'$ then lies in a complementary domain R' of J'. Choose two points A' and B' on J'. There must exist an arc $\langle A'X'B' \rangle$ in R' which has no points in common with M'. For, if every arc in R' joining A' to B' intersects M', there exists a subcontinuum of M' that lies in R' and intersects both arcs of J' which join A' to B'.[11] But this is impossible since T^{-1} preserves regions and since no such subcontinuum of M exists in R.

[10] C. M. Cleveland, *Concerning points of a continuous curve that are not accessible from each other*, Proceedings of the National Academy of Sciences, vol. 13(1927), pp. 275–276.

[11] C. Kuratowski, *Sur les courbes gauches*, Fundamenta Mathematicae, vol. 15(1930), p. 274, Lemma III′.

The arc $\langle A'X'B'\rangle$ lies in a complementary domain D' of M'. We assert that $T(k)$ is contained in the boundary of D'. First, consider any point P on J. If Q is another point of J such that P, Q separate A, B on J, the arc in R' which joins P' to Q' in the complement of M' must meet $\langle A'X'B'\rangle$. Hence P' lies on the boundary of D'. Secondly, consider any point P_0 of k not on J. Since the pairs P_0' and P', P_0' and Q' are known each to lie on a common complementary domain boundary of M', there are arcs in R' and in the complement of M' joining P_0' to P' and P_0' to Q'. Since $P'Q'$ separate $A'B'$, at least one of these arcs must meet $\langle A'X'B'\rangle$. Hence P_0' is in the boundary of D'. Since T^{-1} also preserves regions, it can be similarly shown that if P' is any point in the boundary of D', then $T^{-1}(P')$ is in k. This proves the theorem.

LEMMA 1. *Assume that T preserves sense on the pairs of triods of M. Let U and V_1 represent two non-cut points of the boundary k of a complementary domain D of M, and D' the complementary domain of M' whose boundary is $T(k)$.* (Theorems 10 and 11.) *There exists an arc $\langle\alpha_1\rangle$ in D with end points U and V_1, and an arc $\langle\alpha_1'\rangle$ in D' with ends U' and V_1'. Under the above conditions if T is defined for α_1 and α_1' so that $T(M + \alpha_1) = M' + \alpha_1'$ is a homeomorphism of $M + \alpha_1$, then T and T^{-1} preserve regions in $M + \alpha_1$.*

Proof. It is sufficient to show that T preserves sense on triod pairs in $M + \alpha_1$ (Theorem 10). If U and V_1 are end points in k (i.e., each ϵ-separable from any point in k by a single point), they are also end points in M and no new triods have been introduced by adding α_1 to M. This disposes of the case in which M is acyclic. Hence assume at least one of U, V_1, say V_1, is not an end point of k. Since V_1 is not a cut point of k, it must lie on some closed curve J_1 of k. Assume that U is not on J_1. Choose an arc α in k joining U to V_1. Since k is a c.d.b. of M, every maximal cyclic element of k is a closed curve.[12] But since V_1 is not a cut point, it is a point of Menger order 2 in k. Therefore there exists a subarc α_2 of α with ends U and V_2, where $\alpha_2 \cdot J_1 = V_2$. The two arcs of J_1 joining V_1 to V_2 may each be cut into $\beta_1 + \beta_2$ and $\gamma_1 + \gamma_2$, where $\beta_i \cdot \gamma_i = V_i$ $(i = 1, 2)$. At V_1 and V_2 we then have two triods $[\alpha_1, \beta_1, \gamma_1]$ and $[\alpha_2, \beta_2, \gamma_2]$ which are subsets of a θ-graph, and the sense $\alpha_1\beta_1\gamma_1$ is opposite to the sense $\alpha_2\beta_2\gamma_2$ (Theorem 6). But in like manner in M' the sense $\alpha_1'\beta_1'\gamma_1'$ must be opposite to $\alpha_2'\beta_2'\gamma_2'$, where primed letters correspond under T to the respective unprimed letters. There may exist other triods in $M + \alpha_1$ with vertex at V_1 and including α_1 but not equivalent to τ_1. Suppose such a triod $\tau_3 = [\alpha_1, \beta_3, \gamma_3]$ exists. It can be shown that relative sense on τ_2 and τ_3 is preserved in $M' + \alpha_1'$ by means of a θ-graph containing both τ_2 and τ_3 as follows. Since V_1 is not a cut point of k, it can be shown that β_3 and γ_3 (or equivalent portions) must lie in $\bar{R}$, where R is the complementary domain of k whose boundary is J_1. Since τ_3 is not equivalent to τ_1, there is a point P_1 of τ_3 in R. Suppose P_1 lies on β_3. Let $t = (P_1V_2)$ denote an arc lying except for V_2 in R.

[12] W. L. Ayres, *Continuous curves homeomorphic with the boundary of a plane domain*, Fundamenta Mathematicae, vol. 14(1929), Theorem, p. 92.

Let V_3 denote the first point of β_3 on J_1 from P_1 to V_1. The arc $V_3P_1V_2$ then divides R into two regions r_1 and r_2. Either γ_3, or some subarc γ_4 of γ_3 with V_1 as end point, must lie in $\bar{r}_1$ or $\bar{r}_2$, say $\bar{r}_1$, where γ_1 is in the boundary of r_1. We then obtain a θ-graph containing τ_3 and τ_2, or their equivalents, by joining P_1 to β_2 by an arc in r_2, and joining a point P_2 of γ_4 to γ_2 by an arc in r_1. The three arcs comprising the θ-graph then contain α_1, α_2 ; β_3, β_2 ; and γ_4, γ_2, respectively.

In S' an analogous construction can be made where the arc $V_3'P_1'V_2'$ divides R' into two regions r_1' and r_2', and γ_1' is in the boundary of r_1'. Furthermore, a subarc γ_4' of γ_3' with V_1' as an end point must lie in $\bar{r}_1'$. This is obvious if γ_4' has points in common with γ_1' different from V_1'. If not, one can show that γ_4' lies in r_1' from the fact that sense must be preserved on triod pairs in M. It is possible then in the manner above to obtain a θ-graph such that the three arcs comprising this θ-graph contain α_1', α_2' ; β_3', β_2' ; and γ_4', γ_2' respectively. This demonstrates that τ_2 and τ_3 have opposite sense and in like manner τ_2' and τ_3', as required.

If U is on J_1, then J_1 plus the arc α_1 is a θ-graph, which can be used as before to show that the relative sense of any pair of new triods, with vertices at V_1 and U, is preserved. To complete this case, one then must show that sense is also preserved for some one pair of triods involving one new and one old triod. But if M contains any old triods, there is one such old triod τ_4 with vertex on J_1 and two legs which are subarcs of J_1. If we use the simple closed curve J_1 and a suitable one of the triods τ_1, τ_2, Theorem 7 and the construction of α_1 then show that sense is preserved.

Hence T preserves sense on all triod pairs in $M + \alpha_1$, and therefore T preserves regions in $M + \alpha_1$. In like manner it follows that T^{-1} preserves regions in $M + \alpha_1$.

Proof of Theorem 1. It is possible to carry out this proof by adjoining sufficiently many arcs in the manner of Lemma 1 to extend M to a cyclic curve. The details could be made identical with those used by Gehman (loc. cit., footnote 1). However, we can show that our conditions are equivalent to Gehman's "sides are preserved" as used in the *restricted sense.* We do not prove our condition of sense preserved on triod pairs equivalent to his "sides are preserved under T" in the general form which occurs in his definition, but rather in the more specific and weaker form actually used in the subsequent course of his proof. (See Gehman, loc. cit., p. 263.)

In Lemma 1 let J_1 denote a closed curve composed of the arc α_1 and an arc t in k. Let N_1 denote the subset of M in one complementary domain of J_1, and N_2 the subset of M in the other complementary domain of J_1. Then the result in Lemma 1 shows that $T(N_1)$ lies in one complementary domain of $J_1' = \alpha_1' + T(t)$ while $T(N_2)$ lies in the other complementary domain of J_1'. This is sufficient to show that if T preserves sense on triod pairs of M, then T "preserves sides" in the restricted and weaker form. Therefore T is extendible to S and S'. This completes the proof of Theorem 1.

4. Triods in a Peano space.

THEOREM 12. *If P is a point of Menger order 2 of a triply connected continuum M, then M is locally an arc at P; i.e., there exists a neighborhood of P in which M consists of a simple arc.*

Proof. If M is a closed curve, the theorem is trivial. Hence assume that M is not a closed curve. Let Q represent a point of M that is not of order 2. Such a point Q must exist, otherwise M would be a closed curve.[13] Since P is of order 2, there exists a domain R of M containing P but not Q and such that $F(R)$ consists of two points p and q in M. Since Q is not of order 2, $M - R$ is not an arc. Therefore $\bar{R}$ must be an arc, for otherwise M is not triply connected. This proves the theorem.

THEOREM 13. *The set of all points of order different from 2 in a triply connected continuum M is a closed set.*

This theorem follows directly from Theorem 12.

COROLLARY 1. *If P is a point of order 2 in a triply connected Peano continuum M which is not a simple closed curve, there exists an arc (APB) of M such that the end points A and B are not of order 2 while every point of $\langle APB \rangle$ is of order 2.*

This corollary follows from Theorems 12 and 13.

THEOREM 14. *A point P of a triply connected Peano continuum M is the vertex of a triod in M if and only if P is not of order 2.*

This is an immediate consequence of the "n-Bein" theorem, which states that every point of order at least n is the vertex of an "n-Bein".[14]

LEMMA 2. *If P is an interior point of an arc t in a cyclically connected Peano continuum M, there exists a subarc t' of t which contains P as interior point and which is contained in a simple closed curve J of M.*

Proof. Let A and B represent the end points of t. Since P is not a cut point of M, there exists an arc (AB) of M that does not contain P.[15] The arc t' and the closed curve J may then be obtained as subsets of $t + (AB)$.

COROLLARY 2. *Every triod in a cyclically connected Peano continuum M is equivalent to a subtriod of a θ-graph in M.*

5. **Maximal triply connected curves of M.** The study of triply connected subcontinua of a given space M has some resemblances to the cyclic element theory of Whyburn.

Definition. A triply connected Peano continuum N which is a subset of a

[13] G. T. Whyburn, *On regular points of continua and regular curves of at most order n*, Bulletin of the American Mathematical Society, vol. 35(1929), p. 221, Corollary 2c.

[14] K. Menger, *Kurventheorie*, Leipzig and Berlin, 1932, p. 214.

[15] R. L. Moore, *Concerning continuous curves in the plane*, Mathematische Zeitschrift, vol. 15(1922), p. 255, Theorem 1.

Peano continuum M will be called a *maximal triply connected curve of* M if N is not a proper subset of any other triply connected Peano continuum which is a subset of M.

THEOREM 15. *Every triod in a cyclically connected Peano continuum* M *is equivalent to a subtriod of a maximal triply connected curve of* M.

The following preliminary definitions and lemma are necessary for the proof of Theorem 15.

The topological configuration given by a circle with three radii we shall call a λ-*graph*. It is triply connected. Hence if M contains a subset λ which is a λ-graph, then in any split of M at p, q one component A_λ of $M - (p + q)$ must contain all of λ except perhaps an arc. We shall use the notation $M - (p + q) = A_\lambda + H$, where H then denotes the sum of the remaining components.

Definition. If $M - (p + q) = A_\lambda + H$ is a split at (p, q), this split will be called *minimal with respect to* λ if there is no other split, $M - (p' + q') = A'_\lambda + H'$, in which A'_λ is a proper subset of A_λ.

LEMMA 3. *If* M *is a cyclically connected Peano continuum containing a* λ-*graph, and if* $M - (p' + q') = A'_\lambda + H'$ *is a split, there exists a split* $M - (p + q) = A_\lambda + H$ *which is minimal with respect to* λ *and has* $A_\lambda \subset A'_\lambda$.

Proof. We first study the locus of possible split points p and q. Let U denote any point in H' not in λ. Since M has no cut point, there exists an arc (GUK) of M having its ends G and K on λ. No matter what the position of G and K on λ we can join them by an arc of λ passing through at least three vertices V, V_1, V_2 of λ. This arc added to (GUK) gives a closed curve $J = UV_1VV_2U$. Then any point pair of M which separates U from V must lie on J. In fact one of the points must lie on that subarc (UV_1) of J which does not contain V while the other lies on that subarc (UV_2) which does not contain V.

Let W represent the set composed of all split pairs of M that separate U and V. Then $\overline{W} \subset J$. Let p_0 be the last point of $\overline{W}$ on the arc (UV_1) from U to V_1 and q_0 the last point of $\overline{W}$ on (UV_2) from U to V_2. We shall show that p_0, q_0 is a split pair. Suppose the contrary. There exists then an arc t of M from U to V containing neither p_0 nor q_0. Since $p_0 \subset \overline{W}$, there exists on $(Up_0) \subset (UV_1)$ a point p_1 of a split pair p_1q_1 such that p_1 lies in a neighborhood of p_0 chosen so that (p_1p_0) contains no points of t. Then q_1 lies on the arc (UV_2). In like manner there exists on (Uq_0) a point q_2 of a split pair (p_2q_2) such that the subarc (q_2q_0) of (Uq_0) contains no points of t, while p_2 lies on the arc (UV_1). If q_1 lies between q_2 and q_0 on the arc (UV_2), then p_1q_1 is a split pair not on t, and this is impossible. In like manner p_2 cannot lie between p_1 and p_0 on (UV_1). Hence we have the cyclic order $Up_2p_1p_0Vq_0q_2q_1U$ on J. But p_1 and q_2 are not on t, and t must then contain p_2 and q_1. Suppose q_1 precedes p_2 on t from V to U. Then the subarc (Vq_1) of t plus an arc of J gives an arc of M from V to U that does not contain either of the points of the

split pair p_2, q_2 which were supposed to separate U from V. A similar contradiction is reached if p_2 precedes q_1 on t from V to U. Therefore p_0, q_0 is a split pair, as asserted.

This split at (p_0, q_0) is minimal with respect to λ. For suppose the contrary. Let $M - (p_0 + q_0) = A_\lambda^0 + H^0$ and $M - (p' + q') = A'_\lambda + H'$, where A'_λ is a proper subset of A_λ^0. Since $(p', q') \neq (p_0, q_0)$, at least one of p', q' must lie in A_λ^0. Then p', q' must separate either V and p_0 or V and q_0, and hence p', q' must lie on J and separate V from U. But this violates the choice of p_0 and q_0 above. This proves the lemma, with $(p_0, q_0) = (p, q)$.

Corollary 3. *Neither of the points of the split pair p, q of Lemma 3 is of order 2.*

Proof. Suppose p is of order 2. There exist then two points r and s of M that separate p from U and also p from V. Obviously r and s must lie on the subarc (UpV) of J and one, say r, must lie between p and V on this arc. But then clearly r, q is a split pair violating the choice of the split pair p, q. Hence neither p nor q is of order 2.

Remark. A minimal split found by the above argument is uniquely determined by the point U. Therefore if $M - (p + q) = A_\lambda + H$ and $M - (p' + q') = A'_\lambda + H'$ are any two minimal splits with respect to λ, and if $H \cdot H' \neq 0$, the preceding argument applied to a point U in $H \cdot H'$ shows that these splits are identical, i.e., $H = H'$ and $A_\lambda = A'_\lambda$.

Proof of Theorem 15. *Preliminary remark.* We need a maximal triply connected subset containing a given θ-graph of M. If M is not triply connected, it has a split $M - (p + q) = A_\theta + H$, where A_θ contains at least two of the three independent arcs comprising θ. Since the desired subset cannot be so split, it must consist of a part of A_θ plus at most one arc in H. The first step in obtaining this desired maximal triply connected subset will then replace M by its subset $A_\theta + t$, where t is any one arc of H with ends p and q. This process must in general be applied an infinite number of times.[16]

Any triod of M (M not acyclic) is a subset of a θ-graph θ in M (Corollary 2). Assume θ is not itself a maximal triply connected subset of M. Then it follows without difficulty that θ is a subset of a λ-graph λ of M. Let (p_i, q_i) $(i = 1, 2, 3, \cdots)$ represent the split pairs of M which give minimal splits with respect to λ. These splits are countable. For p_i, q_i can be shown to be local separating points of M, and all save possibly a countable number of the local separating points of any continuum M are points of order 2 in M.[17] But the points p_i, q_i are not of order 2 (Corollary 3).

Let A_i denote the connected component of $M - (p_i + q_i)$ that contains all but perhaps an arc of λ, and let $M_i = A_i + t_i$, where t_i is an arc of M in $M - A_i$

[16] For graphs a similar process appeared in S. Mac Lane, *A structural characterization of planar combinatorial graphs*, this Journal, vol. 3(1937), pp. 460–472.

[17] G. T. Whyburn, *Local separating points of continua*, Monatshefte für Mathematik und Physik, vol. 36(1929), p. 309, Theorem 9.

with ends p_i and q_i. If M does not contain all of λ, we may choose t_i as the remaining arc of λ, so that each M_i contains λ. Let $M_\lambda = M \cdot M_1 \cdot M_2 \cdots$. Then since every M_i is closed, the set M_λ is closed. Since distinct minimal splits give non-intersecting components H_i (see remark following proof of Lemma 3), no two of the arcs t_i intersect except perhaps at end points.

We shall show that M_λ is a maximal triply connected curve of M.

(a) The set M_λ is locally connected. Let A denote any point of M_λ. If A is an interior point of some t_i, then locally M_λ is an arc through A. Hence assume that A is not an interior point of any arc t_i. Since M is locally connected, there exists for any neighborhood R_1 of A a neighborhood R_2 of A such that if B is a point of M_λ in R_2, there exists an arc (BA) of M lying in R_1. If (BA) contains no pair of (p_i, q_i), then (BA) is in M_λ. Assume then that B is not an interior point of any arc t_i and that the pairs (p_a, q_a), (p_b, q_b), (p_c, q_c), $\cdots$ lie on (BA). Let $(BA)' = (BA) \cdot M_\lambda + t_a + t_b + t_c + \cdots$, where t_a is the arc of M_λ with ends p_a, q_a, etc., for t_b, t_c, $\cdots$. At most a finite number of t_a, t_b, t_c, $\cdots$ can be of diameter greater than any ϵ. This follows from the fact that at most a finite number of the sets H_i in $M - (p_i + q_i) = A_i + H_i$ can be of diameter greater than ϵ since M has the Sierpinski property.[18] Hence it can be shown by well-known methods that $(BA)'$ is an arc (of M_λ) from B to A. If each t_a, t_b, t_c, $\cdots$ lies in R_1, then $(BA)'$ is the desired arc. But we could have chosen R_2 so that this is always the case. For there are but a finite number of arcs in the set $\{t_i\}$ of diameter greater than ϵ, hence R_2 can be chosen so that the original arc (BA) of M lies in a neighborhood R_3 of A so small that any t_i with an end p_i or q_i in R_3 different from A lies entirely in the original R_1. If B is an interior point of an arc t_i not ending at A, let q_i denote the first end point of t_i on (AB) from A to B (there may be only one end point of t_i on AB). The modified arc $(AB)'$ in M_λ is then defined as above except that the subarc Bq_i of (BA) is replaced by the subarc Bq_i of t_i. Then $(AB)'$ lies in R_1. If B is an interior point of an arc t_i ending at A, it is sufficient to note that, since at most a finite number of $\{t_i\}$ can have one end outside R_3 while the other end is at A, the choice of R_2 could have been made so that a subarc (BA) of t_i always lies in R_1. Therefore M_λ is a Peano continuum since it is closed and locally connected at any point.[19]

(b) The set M_λ is triply connected. Suppose M_λ is not triply connected, and (p, q) is a split pair, $M_\lambda - (p + q) = H_\lambda + H$, where H_λ is connected (contains all of λ except perhaps an arc) and H is not an arc. We may assume that neither p nor q is an interior point of an arc t_i. For if p (or q) were an interior point of t_i, either end point of t_i taken with q (or p) would yield a split of M_λ. Let A denote a point of H_λ and B any point of H. If (p, q) is not a cut pair

[18] W. Sierpinski, *Sur une condition pour qu'un continu soit une courbe jordanienne*, Fundamenta Mathematicae, vol. 1(1920), pp. 44–60.

[19] It can also be proved that M_λ is a retract of M in the sense of Borsuk. It then follows that M_λ is a Peano continuum by Theorem 7 in Borsuk, *Sur les rétractes*, Fundamenta Mathematicae, vol. 17(1931), pp. 152–170.

of M, there exists an arc (AB) of M that does not contain p or q.[20] Then the corresponding arc $(AB)'$ of M_λ does not contain p or q since neither can be an interior point of an arc of $\{t_i\}$. Assume then that (p, q) is a cut pair of M, but not a split pair. Then $M - (p + q) = G_\lambda + G$, where G_λ is connected and G is an arc which is a subset of H. Let B denote any point in $H - G$. There exists then an arc (AB) of M which does not pass through p or q, and the corresponding arc $(AB)'$ of M cannot pass through p or q. Hence every point of $H - G$ can be connected to the point A through $M_\lambda - (p + q)$ and (p, q) is not a split of M_λ. If (p, q) is a split pair of M, then $H = t_i$ (for some i) since the minimal splits with respect to λ are distinct and H_λ is connected. Hence (p, q) is not a split pair of M_λ, and M_λ must be triply connected.

(c) The set M_λ is a maximal triply connected curve of M. Suppose M_λ is not maximal triply connected. There exists a triply connected Peano continuum $N \subset M$ such that M_λ is a proper subset of N. Hence N is not in every M_i. Assume N not a subset of some M_j. There is a minimal split of M at (p_j, q_j) such that $M - (p_j + q_j) = A_j + H_j$ and $M_j - (p_j + q_j) = A_j + t_j$, where A_j is connected and t_j is an arc in H_j. Then $N - (p_j + q_j) = N \cdot A_j + N \cdot H_j$ gives a split of N. For $N \cdot A_j$ contains all of M except t_j, and $N \cdot H_j$ contains more than t_j since N is not a subset of M_j. But N was assumed triply connected. Therefore M_λ is a maximal triply connected curve of M.

THEOREM 16. *If N is any maximal triply connected curve of a cyclically connected Peano continuum M, then N is either a simple closed curve, or a θ-graph, or one of the curves M_λ constructed above.*

Proof. Assume N is neither a closed curve nor a θ-graph. Then N contains a λ-graph λ. If M is not triply connected, let $M - (p + q) = A_\lambda + H$ represent a minimal split of M. Then since $N - (p + q) = A_\lambda \cdot N + H \cdot N$ is not a split of N, the set $H \cdot N$ is at most an arc t. Hence N is a subset of one of the curves of M_λ. But N is maximal and must then be identical with M_λ.

THEOREM 17. *If N is a maximal triply connected curve of the cyclically connected Peano continuum $M \neq N$, there exists a split pair (p, q) of M such that p and q are the vertices of triods in N.*

Proof. Since $N \neq M$ and N is a maximal triply connected subset of M, N is not a closed curve. If N is a θ-graph, the two vertices of θ must constitute a split pair of M, and these points are vertices of triods in θ. Assume then that $N = M_\lambda$. Let $M - (p + q) = A_\lambda + H$ be a minimal split of M with respect to λ. Let t denote the arc of M_λ in H with ends p and q. There exists a closed curve $J = UpV_1VV_2qU$ of M_λ, where U is an interior point of t, and V_1, V, V_2 are vertices of λ. Assume that at least one of p, q, say p, is of order 2 in M_λ. Then M_λ is locally an arc at p (Theorem 12). Let r denote the first point of order $\neq 2$ in M_λ (Theorem 13) along the subarc (UpV_1) of J, and s the first point of order $\neq 2$ in M_λ along the subarc (UqV_2) of J. Since p is of order 2,

[20] R. L. Moore, loc. cit.

$r \neq p$. Now (r, s) is not a cut pair of M. For then (r, s) would be a split pair with respect to λ. Hence r, s cannot separate U and V in M. There exists then an arc (UV) of M containing neither r nor s. But the corresponding arc $(UV)'$ of M_λ cannot contain either r or s. For if $(UV)'$ contains r (or s), the point r is then an interior point of some arc t_i, since r is not on (UV), and hence is of order 2 in M_λ contrary to assumption. But $(UV)'$ must contain one of (r, s) since obviously (r, s) separate U from V in M_λ. Therefore p, and in like manner q, must be of order $\neq 2$ in M_λ. Then by Theorem 14, p and q are vertices of triods in $M_\lambda = N$.

6. **Proof of Theorem 3, second main theorem.** We shall show that if the conditions for the special list of triods in Theorem 3 are satisfied, then sense is preserved on every triod pair of M.

Let $\tau = [\alpha, \beta, \gamma]$ denote a triod of M with vertex V, where V is not a cut point of M. Since V is the vertex of a triod in M, it is not an end point and hence must be an internal point of some proper cyclic element N of M. Then τ is equivalent to a triod that lies entirely in N. This may be shown as follows: Let t be an arc in $M - V$ joining a point A of α to a point B of β. Let A_1 denote the last point of t on α from A to B and B_1 the first point of t on β. Then the closed curve VA_1B_1V is a subset of N since V is an interior point of N. Hence the subarcs VA_1 of α and VB_1 of β lie in N. In like manner a subarc of γ lies in N. These subarcs constitute a triod equivalent to τ and lying in N. Therefore there exists a maximal triply connected curve M_τ of N containing τ (or an equivalent triod), by Theorem 15. If N is not triply connected, then M_τ contains two triods τ_1 and τ_2 with vertices p and q at a split pair of N (Theorem 17) which is also a split pair of M. But every triply connected Peano continuum has a unique map; i.e., every homeomorphism is extendible in every map of the continuum. Therefore since the condition of Theorem 1 is necessary, sense must be preserved on every triod pair of M_τ, and hence sense is preserved on the triod pairs τ, τ_1 and τ, τ_2. Since the list of Theorem 3 includes either τ_1 or τ_2, it follows then that sense is preserved on every triod pair of N.

If N is a proper subset of M and triply connected, we have our condition on one triod of N, and since N has a unique map, the condition is satisfied on every triod pair of N. Therefore sense is preserved on every triod pair of M and T can be extended to S and S' (Theorem 1).

The conditions are obviously necessary since the condition in Theorem 1 is necessary.

UNIVERSITY OF ARKANSAS AND HARVARD UNIVERSITY.

A GENERALIZATION OF POISSON'S SUMMATION FORMULA

By S. Bochner

Poisson's formula. The standard form of Poisson's formula is[1]

$$\sum_{-\infty}^{\infty} f(m) = \sum_{-\infty}^{\infty} F(2\pi n), \tag{1}$$

$$F(\alpha) = \int_{-\infty}^{\infty} e^{-i\alpha x} f(x)\, dx. \tag{2}$$

If $f(z)$ is analytic in a strip

$$|\, y \,| < y_0 \tag{3}$$

of the complex plane $z = x + iy$, $\sum f(m)$ is the sum of the residues of the function

$$\pi f(z) \frac{\cos \pi z}{\sin \pi z} \tag{4}$$

and therefore it is the limit, as $T \to \infty$, of the Cauchy integral of the function (4) around the rectangle with the corners $\pm T \pm bi$ $(b < y_0)$. In order to transform the integral into $\sum F(2\pi n)$ we have to replace $-i \cot \pi z$ by the expansion

$$1 + 2 \sum_{n=1}^{\infty} e^{-2n\pi i z} \tag{5}$$

for $y < 0$ and by

$$-(1 + 2 \sum_{n=1}^{\infty} e^{2n\pi i z}) \tag{6}$$

for $y > 0$.[2]

In our generalizations we will take an unspecified meromorphic function $\varphi(z)$ in a strip (3) instead of the particular function $\cot \pi z$. This will lead to a formula

$$\sum_{-\infty}^{\infty} r_m f(a_m) = \int_{-\infty}^{\infty} F(\alpha)\, d\Phi(\alpha). \tag{7}$$

The numbers a_m will be simple poles of $\varphi(z)$ and r_m their residues, and the weight function $\Phi(\alpha)$ will be taken from general expansions, analogous to (5) and (6), of the function $\varphi(z)$ in two strips in which it has no poles.

Received October 18, 1939.

[1] Compare S. Bochner, *Fouriersche Integrale*, p. 33; E. C. Titchmarsh, *Fourier Integrals*, p. 60.

[2] Compare E. Lindelöf, *Calcul des Résidus*, Chapter III.

More general than (7) is the formula

$$\int_{-\infty}^{\infty} f(a)\,dr(a) = \int_{-\infty}^{\infty} F(\alpha)\,d\Phi(\alpha). \tag{8}$$

It arises formally by using a function $\varphi(z)$ which is not necessarily meromorphic but has a representation of the type

$$\varphi(z) = \int_{-\infty}^{\infty} \frac{dr(a)}{z-a}, \tag{9}$$

or which, more generally, in rectangles with the corners $\pm T \pm bi$, for a sequence of increasing values T, differs from

$$\int_{-T}^{T} \frac{dr(a)}{z-a}$$

by functions which are analytic in the rectangles. Now, formula (8) has a very familiar appearance. It is Parseval's formula for Fourier integrals, the "differential" $d\Phi(\alpha)$ being the Fourier transform of the differential $dr(a)$. As a matter of fact, formula (1) is a very particular case of (8) insofar as in (1) the differential $dr(a)$ of (8) is essentially self-reciprocal. We will not investigate the validity of the more general formula (8) although our theorems could be easily extended to cover it in the main.

The generalized formula. In our theorems the assumptions will be somewhat elaborate. However, they will refer mostly to the underlying meromorphic function $\varphi(z)$ rather than to the function $f(z)$ which is to be summed.

THEOREM 1. *Assumptions*: (i) *The function $\varphi(z)$ is defined and meromorphic in the strip* (3). *All poles of $\varphi(z)$ in* (3) *are simple and they are located in a closed interior strip*

$$|y| \leqq \delta \qquad (0 \leqq \delta < y_0). \tag{10}$$

The set of all poles will be denoted by $\{a_m\}$, the corresponding residues by $\{r_m\}$.

(ii) *There exists an unbounded set of positive abscissas $x = T$ such that $\varphi(z)$ is bounded on the totality of segments*

$$x = \pm T, \qquad |y| < y_0. \tag{11}$$

(iii) *$\varphi(z)$ can be represented by an absolutely convergent Stieltjes-Laplace integral*

$$\varphi(z) = 2\pi i \int_{-\infty}^{\infty} e^{-i\alpha z}\,d\Phi_1(\alpha) \tag{12}$$

in the open border strip $\delta < y < y_0$, and by a similar integral

$$\varphi(z) = 2\pi i \int_{-\infty}^{\infty} e^{-i\alpha z}\,d\Phi_2(\alpha) \tag{13}$$

in the border strip $-y_0 < y < -\delta$. *We shall write*

$$\Phi(\alpha) = \Phi_2(\alpha) - \Phi_1(\alpha).$$

(iv) $f(z)$ *is analytic in* (3), *and tends to* 0 *as* $x \to \pm\infty$ *and for each* y *in* (3) *the limit*

$$e^{-\alpha y}F(\alpha) = \lim \int_{-T}^{T} e^{-i\alpha x} f(x + iy)\, dx \qquad (T \to \infty) \tag{14}$$

exists boundedly in all α *and uniformly in every finite* α*-interval.*

Conclusion:

$$\lim_{T\to\infty} \sum_{-T}^{T} r_m f(a_m) = \int_{-\infty}^{\infty} F(\alpha)\, d\Phi(\alpha), \tag{15}$$

the sum extending over all zeros a_m *whose real part lies between* $-T$ *and* T. *The integral converges absolutely.*

Remark about Theorem 1. Assumption (iii) implies that $\varphi(z)$ is bounded in every closed strip in the interior of the two border strips. We observe that boundedness alone[3] would allow us to write down the integrals (12) and (13) with sufficiently general transforms $\Phi_1(\alpha)$ and $\Phi_2(\alpha)$ and as a matter of fact we might continue the discussion on that basis. However, we prefer to avoid complications and so we explicitly assume that the transforms are of bounded variation in every finite interval and that the integrals are absolutely convergent in their respective strips.

As for $f(z)$ we observe that assumption (iv) is certainly fulfilled if $f(z)$ belongs to the Lebesgue class L_1 in the strip (3).

Proof of Theorem 1. Applying Cauchy's theorem to the function $f(z)\cdot\varphi(z)$ around the rectangle with the vertices $\pm T \pm ib$, we easily obtain

$$\lim_{T\to\infty} \left\{ \sum_{-T}^{T} r_m f(a_m) - \frac{1}{2\pi i} \int_{-T}^{T} (f(x - ib)\varphi(x - ib) - f(x + ib)\varphi(x + ib))\, dx \right\} = 0. \tag{16}$$

Now

$$\int_{-T}^{T} f(x + ib)\varphi(x + ib)\, dx = 2\pi i \int_{-\infty}^{\infty} \left(\int_{-T}^{T} f(x + ib) e^{-i\alpha(x+ib)}\, dx \right) d\Phi_1(\alpha)$$

and by assumptions (iii) and (iv) this tends to $\int_{-\infty}^{\infty} F(\alpha)\, d\Phi_1(\alpha)$ as $T \to \infty$. A similar limit relation holds for $-b$ instead of b, and this proves (15).

[3] Compare S. Bochner, *Verallgemeinerte Fourier und Laplace-Integrale*, Math. Annalen, vol. 97(1927), pp. 635–662.

Almost-periodic cases. If $\Phi(\alpha)$ is a pure jump function having jumps γ_n at the points $\alpha = \alpha_n$, we can write

$$\lim_{T\to\infty} \sum_{-T}^{T} r_m f(a_m) = \sum_{-\infty}^{\infty} \gamma_n F(\alpha_n). \tag{17}$$

To assume that $\Phi_1(\alpha)$ and $\Phi_2(\alpha)$ are jump functions separately is equivalent to assuming that $\varphi(z)$ is almost periodic in each of the two border strips and that the corresponding Dirichlet expansions are absolutely convergent.[4] In this connection it is interesting to write our formula replacing $f(z)$ by $f(z + t)$ in the seemingly more general form

$$\lim_{T\to\infty} \sum_{-T+t}^{T+t} r_m f(a_m + t) = \sum_{-\infty}^{\infty} \gamma_n F(\alpha_n) e^{i\alpha_n t}, \qquad -\infty < t < \infty. \tag{18}$$

This formula shows that the poles a_m if counted with their "multiplicities" r_m are distributed in such a way that the expression on the left represents an almost-periodic function of the real variable t for any admissible function $f(z)$. Introducing the function $r(a)$ as in (8), we see that the left side of (18) is formally

$$\int_{-\infty}^{\infty} f(a + t)\, dr(a) \equiv \int_{-\infty}^{\infty} f(a)\, d_a r(a - t),$$

and we may express our last conclusion by stating that the "differential" $dr(a)$ is a weakly almost-periodic function of a.

Moreover, we may drop the assumption that the two Dirichlet series of $\varphi(z)$ are absolutely convergent. Approximating to $\varphi(x + ib)$ and $\varphi(x - ib)$ by finite exponential sums, we easily conclude that the left side of (18) is an almost-periodic function in t and that the right side is its formal Fourier expansion. A peculiar case arises if

$$\varphi(z) = \frac{\psi'(z)}{\psi(z)},$$

the function $\psi(z)$ being almost periodic in (3) and $\neq 0$ in the border strips; the poles a_m of $\varphi(z)$ are the zeros of $\psi(z)$ and the residues are all $+1$. In this case relation (13) illustrates known facts about the distribution of such zeros.[5]

Other cases. For typical classes of functions $f(z)$ the limit (14) breaks down for $\alpha = 0$. We are therefore going to formulate a theorem which will take care of that possibility.

THEOREM 2. *Assumptions*: (i), (ii), (iii) *as in Theorem* 1.

(iv) *In some interval* $|\alpha| < \alpha_1$ *the functions* $\Phi_1(\alpha)$ *and* $\Phi_2(\alpha)$ *and hence the function* $\Phi(\alpha)$ *itself are constants except for jumps at* $\alpha = 0$. *The jump of* $\Phi(\alpha)$ *will be denoted by c, and the function resulting from* $\Phi(\alpha)$ *by omission of that jump will be denoted by* $\Psi(\alpha)$.

[4] A. S. Besicovitch, *Almost Periodic Functions*, Chapter III.

[5] B. Jessen, *Über die Nullstellen einer analytischen fastperiodischen Funktion*, Math. Annalen, vol. 108(1933), pp. 485–516.

(v) *The function $f(z)$ is analytic in* (3) *and tends to* 0 *as $x \to \pm\infty$, and for each y in* (3) *the limit* (14) *exists boundedly in the half-lines $|\alpha| \geqq \alpha_0$ for each $\alpha_0 > 0$ and uniformly in every finite interval of the half-lines.*

Conclusion:

$$\lim_{T\to\infty}\left(\sum_{-T}^{T} r_m f(a_m) - c\int_{-T}^{T} f(x)\,dx\right) = \int_{-\infty}^{\infty} F(\alpha)\,d\Psi(\alpha), \tag{19}$$

the integral converging absolutely.

Proof. Denoting the jumps of $\Phi_1(\alpha)$ and $\Phi_2(\alpha)$ at the origin by $-c_1$ and c_2 and putting

$$\psi_1(x+ib) = 2\pi i(\varphi(x+ib)+c_1) = \int_{-\infty}^{\infty} e^{-i\alpha(x+ib)}\,d\Psi_1(\alpha),$$

$$\psi_2(x-ib) = 2\pi i(\varphi(x-ib)-c_2) = \int_{-\infty}^{\infty} e^{-i\alpha(x-ib)}\,d\Psi_2(\alpha),$$

we have $c = c_1 + c_2$, $\Psi(\alpha) = \Psi_2(\alpha) - \Psi_1(\alpha)$ and relation (19) can be derived from (16).

Finally we want to consider the case in which $f(z)$ is defined only in the closed strip $|y| \leqq \delta$ and not in a larger strip. (We recall that in defining the quantity δ we admitted the possibility $\delta = 0$.)

THEOREM 3. *Assumptions*: (i), (ii), (iii) *as in Theorem* 1.

(iv) *The sum of the absolute values $|r_m|$ extending over those poles a_m whose real part lies between t and $t+1$ is bounded in $-\infty < t < \infty$.*

(v) *The function $f(z)$ is defined and measurable in $|y| \leqq \delta$ and for $\delta > 0$ it is analytic in $|y| < \delta$. The function $g(t)$ which is the least upper bound of $|f(x+yi)|$ for $|y| \leqq \delta$ and $t < x < t+1$ belongs to class L_1 in $-\infty < t < \infty$.*

Conclusion:

$$\sum_{-\infty}^{\infty} r_m f(a_m) = \lim_{A\to\infty}\int_{-A}^{A}\left(1 - \frac{|\alpha|}{A}\right)F(\alpha)\,d\Phi(\alpha), \tag{20}$$

the sum converging absolutely.

Proof. By (v), $f(z)$ has a transform $F(\alpha)$ as in previous cases, and for each finite A the approximating function

$$f_A(z) = \frac{1}{2\pi i}\int_{-A}^{A}\left(1 - \frac{|\alpha|}{A}\right)e^{i\alpha z}F(\alpha)\,d\alpha$$

belongs to class L_1 for all y. Thus, by Theorem 1,

$$\lim_{T\to\infty}\sum_{-T}^{T} r_m f_A(a_m) = \int_{-A}^{A}\left(1 - \frac{|\alpha|}{A}\right)F(\alpha)\,d\Phi(\alpha).$$

If we let $A \to \infty$, the assumptions (iv) and (v) justify an interchange of limits[6] and we finally obtain (20).

[6] N. Wiener, *Tauberian theorems*, Annals of Math., vol. 33(1932), in particular pp. 21–24.

Constructing $f(z)$ in Theorem 3 appropriately, we obtain the following result.

Corollary. *If a function $\varphi(z)$ satisfies all assumptions of Theorem 3, then the integrals*

$$\int_1^\infty e^{-\alpha\delta} M(\alpha)\, d\Phi(\alpha), \qquad \int_1^\infty e^{-\alpha\delta} M(\alpha)\, d\Phi(-\alpha)$$

converge $(C, 1)$ for any decreasing convex multiplier $M(\alpha)$ for which

$$\int_1^\infty M(\alpha)\, d\alpha < \infty.$$

For example, the standard functions α^{-p-1}, $\alpha^{-1} (\log \alpha)^{-p-1}$, etc. $(p > 0)$ are admissible multipliers.

A remark on Dirichlet series. As an illustration of the last result consider a function $\varphi(z)$ which is meromorphic in the whole plane. All its poles $\{z_m\}$ are simple and lie in a strip $-\delta \leqq x \leqq \delta$ and there exists a number $\delta_1 > \delta$ and a sequence $Y_k \to \infty$ such that $\varphi(z)$ is bounded on the totality of segments $-\delta_1 \leqq x \leqq \delta_1$, $y = \pm Y_k$. Also if r_m denotes the residue at z_m, then the sum of $| r_m |$ extending over those z_m whose imaginary part lies between y and $y + 1$ is bounded in y. Furthermore, in each of the half-planes $x > \delta$ and $x < -\delta$, $\varphi(z)$ is an absolutely convergent Dirichlet series, and let the series in the right half-plane be denoted by

$$\sum \gamma_n e^{-\alpha_n z} \qquad (\alpha_n \geqq 0).$$

Then, as a consequence of the corollary,

$$\lim_{A\to\infty} \sum_{1 \leqq \alpha_n \leqq A} \left(1 - \frac{|\alpha_n|}{A}\right)^\sigma M(\alpha_n) \gamma_n e^{-\alpha_n z}$$

exists for $\sigma = 1$ for every point of the line $x = \delta$ itself. As a matter of fact, the result also holds if the Dirichlet series are not absolutely convergent but only bounded in every interior half-plane, and for any $\sigma > 0$. However, for Dirichlet series with positive coefficients γ_n the theorem of Landau and Ikehara[7] is a much stronger statement.

Princeton University.

[7] S. Bochner, *Ein Satz von Landau und Ikehara*, Math. Zeitschrift, vol. 37 (1933), pp. 1–9; E. Landau, *Über Dirichletsche Reihen*, Göttinger Nachrichten, 1932, pp. 525–527.

BERTRAND CURVES AND HELICES

By James K. Whittemore

Introduction. Every student of classical differential geometry meets early in his course the subject of Bertrand curves, discovered in 1850 by J. Bertrand.[1] A Bertrand curve is a curve such that its principal normals are the principal normals of a second curve. It is proved in most texts on the subject that the characteristic property of such a curve is the existence of a linear relation between the curvature and the torsion; the discussion appears as an application of the Frenet-Serret formulas.[2] The paper here presented originated in an attempt to specify exactly what Bertrand curves exist having a given linear relation between curvature and torsion, and how such curves may be found. It is certainly well known that a curve is determined by its curvature and torsion uniquely except as to its position in space, more precisely that a curve whose curvature and torsion are given functions of its arc is found by the integration of a Riccati equation.[3] It is, however, impossible to integrate the equation for given curvature and torsion except in some simple cases, usually such that the required curve is a helix. The general theorem of Lie is of interest and value, but it does not give very definite information about the Bertrand curves. J. A. Serret proved in 1850 that curves of given constant curvature and curves of given constant torsion can be found by quadratures. These two kinds of curves and helices are all in a sense Bertrand curves arising in exceptional cases of the linear relation between curvature and torsion, as will appear below. L. Bianchi proved[4] that Bertrand curves for a given linear relation can be found by quadratures, so that in these respects this paper has nothing to add. It is, however, here proved that we may find by quadratures a unique Bertrand curve for a given linear relation with an arbitrary spherical representation; that the coördinates of any point of this curve are expressed in a simple way in terms of the coördinates of certain curves of constant curvature and of constant torsion both of which have the same spherical representation as the Bertrand curve. The study and use of the spherical representation suggests the determination of the spherical representation of the helices of a sphere. This discussion is carried out in the second section of this paper and the spherical helices found from their spherical representation. It is shown that the projection of

Received October 27, 1939.

[1] For this and other historical references see Encyclopädie der Mathematischen Wissenschaften, vol. 3, D1,2, p. 82ff.

[2] E.g., Picard, *Traité d'Analyse*, 2d ed., vol. 1, p. 394.

[3] Sophus Lie, 1882.

[4] *Geometria Differenziale*, p. 31. See also Darboux, *Théorie Générale des Surfaces*, vol. 1, pp. 42–45.

every spherical helix on a plane perpendicular to its axis is an epicycloid. The remainder of the paper is devoted to the study of helices on surfaces of revolution of the second degree, in each case a surface of revolution of a conic about one of its principal axes and, in all cases, the axis of the helix coinciding with that of the surface. In the third section it is shown that the helices on any surface of revolution, axes of curve and surface the same, are given by a quadrature, an integration which can be carried out for the surfaces of the second degree. The helices of the paraboloid and some of the helices of the hyperboloids are determined by the method indicated. It is proved that the projections of the parabolic helices on a plane perpendicular to the axis are involutes of circles. In the fourth and last section of the paper a method is found for determining surfaces of revolution carrying helices with a given projection on a plane perpendicular to the axis. By this method, some of the preceding results are verified, and the helices of the hyperboloids, not found in the third section, and those of the ellipsoid are obtained. It appears that the projections of all ellipsoidal helices are epicycloids, those of the hyperbolic helices of this section, hypocycloids.

1. **Spherical representation. Bertrand curves.** We use the notation of Eisenhart's *Differential Geometry*. The coördinates of a point of a curve x, y, z are considered as functions of the arc of the curve s and we assume that for the part of the curve discussed the third derivatives of the coördinates with respect to the arc exist. In general

$$\left(\frac{dx}{ds}\right)^2 + \left(\frac{dy}{ds}\right)^2 + \left(\frac{dz}{ds}\right)^2 = 1.$$

The direction cosines of the positive direction of the tangent are α, β, γ; either direction of the tangent may be chosen as the positive direction and the arc is measured in the direction of the positive tangent so that $\alpha = dx/ds$ with similar expressions for β and γ. The direction cosines of the positive direction of the principal normal, the direction from the point of the curve to the corresponding center of curvature, are l, m, n. The direction cosines of the positive binormal are λ, μ, ν; the positive direction of the binormal is chosen so that

$$\begin{vmatrix} \alpha & \beta & \gamma \\ l & m & n \\ \lambda & \mu & \nu \end{vmatrix} = 1;$$

that is, so that the positive tangent, principal normal and binormal have the disposition of the x-, y-, z-axes. The curvature is $1/\rho$ and is positive and the torsion is $1/\tau$. With this notation the Frenet-Serret formulas for the derivatives with respect to s of the nine direction cosines are

$$\frac{d\alpha}{ds} = \frac{l}{\rho}, \qquad \frac{dl}{ds} = -\frac{\alpha}{\rho} - \frac{\lambda}{\tau}, \qquad \frac{d\lambda}{ds} = \frac{l}{\tau}$$

with six others obtained from these by advancing the letters α, l, λ.

The spherical representation or indicatrix of a curve C is a curve C_1 on the unit sphere given by

$$x_1 = \alpha, \qquad y_1 = \beta, \qquad z_1 = \gamma.$$

We use for C_1 the same letters as for C with the subscript 1 so that, for example, the arc of C_1 is denoted by s_1. Then

$$\alpha_1 = \frac{dx_1}{ds_1} = \frac{d\alpha}{ds}\frac{ds}{ds_1} = \frac{l}{\rho}\frac{ds}{ds_1}.$$

Squaring and adding the three similar equations, we get

$$1 = \frac{1}{\rho^2}\left(\frac{ds}{ds_1}\right)^2, \qquad \rho ds_1 = ds,$$

if we agree to choose the positive direction of the tangent to C_1 to correspond to that of C; with this convention $\alpha_1 = l$.

We express the determinants Δ and Δ', given in terms of the elements of C_1, in terms of the elements of C:

$$\Delta = \begin{vmatrix} x_1 & y_1 & z_1 \\ \dfrac{dx_1}{ds_1} & \dfrac{dy_1}{ds_1} & \dfrac{dz_1}{ds_1} \\ \dfrac{d^2x_1}{ds_1^2} & \dfrac{d^2y_1}{ds_1^2} & \dfrac{d^2z_1}{ds_1^2} \end{vmatrix} = -\frac{\rho}{\tau}, \qquad \Delta' = \begin{vmatrix} \dfrac{dx_1}{ds_1} & \dfrac{dy_1}{ds_1} & \dfrac{dz_1}{ds_1} \\ \dfrac{d^2x_1}{ds_1^2} & \dfrac{d^2y_1}{ds_1^2} & \dfrac{d^2z_1}{ds_1^2} \\ \dfrac{d^3x_1}{ds_1^3} & \dfrac{d^3y_1}{ds_1^3} & \dfrac{d^3z_1}{ds_1^3} \end{vmatrix} = -\rho\frac{d}{ds}\left(\frac{\rho}{\tau}\right) = \frac{d\Delta}{ds_1}.$$

The values of Δ and Δ' as given are easily found if we use the Frenet formulas with the value of the determinant $|\alpha\ m\ \nu| = 1$. The value of Δ is important for the discussion following. It is interesting to note that for any curve the ratio ρ/τ is expressed in terms of the elements of the spherical indicatrix. It is curious that $\Delta' = d\Delta/ds_1$ although Δ' as given is not the formal derivative of Δ with respect to s_1.

The general linear relation of curvature and torsion for a curve is

$$\frac{A}{\rho} + \frac{B}{\tau} = C,$$

where A, B, C are constants not all zero. The exceptional cases, previously referred to, are

1. $A = 0$. The curve has constant torsion.
2. $B = 0$. The curve has constant curvature.
3. $C = 0$. The ratio ρ/τ is constant and the curve is a helix.

These three types of curves we shall not hereafter call Bertrand curves. For a Bertrand curve then no one of the constants A, B or C is zero and the relation can be written

$$\frac{a}{\rho} + \frac{b}{\tau} = 1,$$

where a and b are both constants different from zero.

Since for any curve $\alpha = dx/ds$, we have

$$x = \int \alpha \, ds = \int x_1 \rho \, ds_1,$$

with similar expressions for y and z. If ρ is a given positive function of s, we may regard it as a function of s_1 for $ds = \rho \, ds_1$, or if ρ is any positive function of s_1, and if x_1, y_1, z_1 are any three functions of x_1 such that $x_1^2 + y_1^2 + z_1^2 = 1$, then the three integrals give the coördinates of a curve C having the spherical representation $C_1(x_1, y_1, z_1)$ and curvature $1/\rho$.

Any curve of constant curvature, $\rho = a > 0$, is given by $x = a \int x_1 \, ds_1$ with similar equations for y and z, where $C_1(x_1, y_1, z_1)$ is an arbitrary curve of the unit sphere. If we now think of C_1 as a given curve of the unit sphere, more precisely of x_1, y_1, z_1 as any three functions of s_1 such that

$$x_1^2 + y_1^2 + z_1^2 = x_1'^2 + y_1'^2 + z_1'^2 = 1,$$

then Δ is a known function of s_1. For any curve of constant torsion, $\tau = b \gtrless 0$, with the given spherical representation, $\rho = -b\Delta$ and $x = -b \int x_1 \Delta \, ds_1$ with similar equations for y and z. For a Bertrand curve with the given spherical representation such that

$$\frac{a}{\rho} + \frac{b}{\tau} = 1$$

we have $\rho = a - b\Delta$ and

$$x = a \int x_1 \, ds_1 - b \int x_1 \Delta \, ds_1$$

with similar equations for y and z. The result is this: The coördinates of a Bertrand curve with the linear relation above and an arbitrary spherical representation are, if a is positive, given by the sum of the corresponding coördinates of the curves of constant curvature, $\rho = a$, and of constant torsion, $\tau = b$, which both have the same spherical representation. For negative a this statement requires a slight modification.

As to arbitrary spherical representation, the question may be put in this form: how can x_1, y_1, z_1 be chosen as functions of s_1 so that

$$x_1^2 + y_1^2 + z_1^2 = 1, \qquad x_1'^2 + y_1'^2 + z_1'^2 = 1?$$

The first of these equations may be replaced by

$$x_1 = \cos u \cos v, \qquad y_1 = \cos u \sin v, \qquad z_1 = \sin u.$$

Substituting in the second equation, we get

$$u'^2 + v'^2 \cos^2 u = 1,$$

where u' and v' are the derivatives of u and v, respectively, with respect to s_1. Then u may be chosen arbitrarily as a function of s_1 and v is given by a quadrature.

2. **The spherical helix.** It is of interest to consider the helix from the point of view of its spherical representation. There are two definitions of a helix which are well known to be equivalent: first, a helix is a curve such that the ratio ρ/τ is constant; secondly, a helix is a curve whose tangent makes a constant angle with a fixed direction. The equivalence of the definitions is clearly exhibited by the equation

$$\Delta' = -\rho \frac{d}{ds}\left(\frac{\rho}{\tau}\right).$$

It is, in fact, only for the purpose of showing this equivalence that this equation is given.

Consider now a helix whose tangent makes a constant angle $C \neq \frac{1}{2}\pi$ with the z-axis; we may without restriction suppose C acute. We have

$$\gamma = \cos C, \qquad \frac{d\gamma}{ds} = \frac{n}{\rho} = 0, \qquad n = 0.$$

From $\gamma^2 + n^2 + \nu^2 = 1$ follows $\nu = \pm \sin C$. From

$$\frac{dn}{ds} = -\frac{\gamma}{\rho} - \frac{\nu}{\tau} = 0$$

follows

$$\frac{\rho}{\tau} = -\frac{\gamma}{\nu} = \mp \cot C.$$

Assuming the torsion of the helix to be positive, we see that

$$\nu = -\sin C, \qquad \frac{\rho}{\tau} = \cot C.$$

If now we consider a helix on a sphere of radius R, we have also the equation[5]

$$\rho^2 + \tau^2 \left(\frac{d\rho}{ds}\right)^2 = R^2,$$

[5] Eisenhart, *Differential Geometry*, pp. 36, 37.

an equation true for all curves on the sphere except the circles. If we take $R = 1$ for simplicity, and replace τ by $\rho \tan C$ and ds by $\rho\, ds_1$, the last equation becomes

$$\rho^2 + \tan^2 C \left(\frac{d\rho}{ds_1}\right)^2 = 1$$

and

$$\text{arc} \sin \rho = \pm s_1 \cot C + \text{constant}.$$

We assume the helix to pass through the point (0, 1, 0) and suppose s measured from this point in the direction of increasing z. Since at this point the principal normal of the helix is normal to the sphere, it follows from Meusnier's theorem[6] that there $\rho = 1$, and if we choose s_1 also zero for the point, the constant of the last equation is determined as $\frac{1}{2}\pi$; then $\rho = \cos (s_1 \cot C)$. To find the spherical representation of the helix we have

$$x_1^2 + y_1^2 + z_1^2 = 1, \qquad z_1 = \cos C.$$

We write $x_1 = \sin C \cos u$, $y_1 = \sin C \sin u$ and substitute x_1, y_1, z_1 in the equation $\Delta = -\rho/\tau$ obtaining $u = -s_1 \csc C$, so that

$$x_1 = \sin C \cos (s_1 \csc C), \qquad y_1 = -\sin C \sin (s_1 \csc C), \qquad z_1 = \cos C.$$

The equations of the helix are now given by

$$\begin{aligned} x &= \int \rho x_1 \, ds_1 = \int \cos (s_1 \cot C) \sin C \cos (s_1 \csc C)\, ds_1, \\ y &= \qquad\qquad -\int \cos (s_1 \cot C) \sin C \sin (s_1 \csc C)\, ds_1, \\ z &= \qquad\qquad \int \cos (s_1 \cot C) \cos C \, ds_1 . \end{aligned}$$

Carrying out these integrations and determining the constants so that for $s_1 = 0$ the point of the helix is (0, 1, 0), we have

$$\begin{aligned} x &= \sin^2 \tfrac{1}{2}C \sin (s_1 \cot \tfrac{1}{2}C) + \cos^2 \tfrac{1}{2}C \sin (s_1 \tan \tfrac{1}{2}C), \\ y &= \sin^2 \tfrac{1}{2}C \cos (s_1 \cot \tfrac{1}{2}C) + \cos^2 \tfrac{1}{2}C \cos (s_1 \tan \tfrac{1}{2}C), \\ z &= \sin C \sin (s_1 \cot C). \end{aligned}$$

If we consider in its entirety the curve given by these equations, it appears that the maximum z is equal to $\sin C$, given by $s_1 = \frac{1}{2}\pi \tan C$; for this point of the helix we have $dx = dy = dz = 0$; the point is actually a cusp. The length of the helix from $z = 0$ to the cusp is $\tan C$. Beyond the cusp the torsion becomes negative.

[6] Eisenhart, loc. cit., p. 118.

The projection of the helix on the xy-plane, a plane perpendicular to the axis, is given by the first two of the preceding group of equations. We write in these

$$\cos^2\frac{C}{2} = R + r, \qquad \sin^2\frac{C}{2} = r, \qquad s_1 \tan\frac{C}{2} = \theta.$$

From this

$$R = \cos C, \qquad R + 2r = 1, \qquad \frac{R + r}{r}\,\theta = s_1 \cot\frac{C}{2}.$$

Then the equations of the projection of the helix are

$$x = (R + r)\sin\theta + r\sin\frac{R + r}{r}\,\theta,$$

$$y = (R + r)\cos\theta + r\cos\frac{R + r}{r}\,\theta.$$

These are the equations of an epicycloid though not quite the "standard" equations given in elementary texts, for $\theta = 0$ gives the point $(0, R + 2r)$ and the cusps are given by values of θ equal to odd multiples of $r\pi/R$.

3. **Helices on surfaces of revolution.** It appears to be impossible to determine helices on other surfaces by finding their spherical representations; for on no surface other than the sphere is there a relation between curvature and torsion holding for all curves of the surface, but the problem of finding the helices on a surface may be stated directly and, in some cases, solved. We can immediately write a differential equation for the helices of any given surface having as axis the z-axis.

If a surface is given in terms of parameters u, v, and if for the surface the linear element is $ds^2 = E\,du^2 + 2F\,du\,dv + G\,dv^2$, the differential equation of helices making with the z-axis a fixed acute angle C is

$$\frac{dz}{ds} = \frac{z_u\,du + z_v\,dv}{[E\,du^2 + 2F\,du\,dv + G\,dv^2]^{\frac{1}{2}}} = \cos C,$$

where z_u and z_v are the partial derivatives of z with respect to the parameters. When the given surface is one of revolution, whose axis is also the z-axis, so that the surface and the helix have the same axis, the integration of the differential equation can be reduced to a quadrature, a quadrature which can be carried out for surfaces of the second degree. In this and the next sections are found the helices for all cases where the meridian of the surface is a conic and the common axis of the surface and the helix is a principal axis of the conic. In this section we find first the helices of the paraboloid and show that their projections on a plane perpendicular to the axis are involutes of circles; secondly, we find some of the helices of the hyperboloids. In the fourth and last section the other cases named above are discussed by a different method.

For a surface of revolution whose axis is the z-axis

$$x = u \cos v, \qquad y = u \sin v, \qquad z = f(u),$$

where $z = f(u)$ is the equation of the meridian. For this case

$$E = 1 + f'^2, \qquad F = 0, \qquad G = u^2.$$

The integration of the differential equation above is given by

$$v = \pm \int [f'^2 \tan^2 C - 1]^{\frac{1}{2}} \frac{du}{u} + K.$$

To simplify the discussion we choose K equal to zero, other values of K obviously giving the same helix revolved about the z-axis; also we use only the positive sign before the integral, change of sign obviously giving a helix symmetrical to the first with respect to the yz-plane.

We consider first the paraboloid, meridian $u^2 = 2pz$. Substituting $f' = u/p$, we obtain

$$\begin{aligned} v &= \frac{\tan C}{p} \int [u^2 - p^2 \cot^2 C]^{\frac{1}{2}} \frac{du}{u} \\ &= \frac{\tan C}{p} \left\{ [u^2 - p^2 \cot^2 C]^{\frac{1}{2}} - p \cot C \operatorname{arc} \cos \left(\frac{p \cot C}{u} \right) \right\}. \end{aligned}$$

Setting

$$\varphi = \operatorname{arc} \cos \left(\frac{p \cot C}{u} \right), \qquad \theta = \tan \varphi, \qquad R = p \cot C,$$

we find

$$x = u \cos v = R \cos \theta + R\theta \sin \theta, \qquad y = u \sin v = R \sin \theta - R\theta \cos \theta;$$

this shows that the projection of the helix on the xy-plane is an involute of a circle.

Consider next helices on the hyperboloid of revolution of one sheet, meridian $u^2 - z^2/b^2 = 1$, supposing $b \tan C > 1$. We have

$$f(u) = b\,[u^2 - 1]^{\frac{1}{2}}, \qquad f' = \frac{bu}{[u^2 - 1]^{\frac{1}{2}}},$$

$$v = \int \left[\frac{u^2(b^2 \tan^2 C - 1) + 1}{u^2 - 1} \right]^{\frac{1}{2}} \frac{du}{u}.$$

We write

$$[b^2 \tan^2 C - 1]^{\frac{1}{2}} = \frac{1}{a} = \cot \alpha;$$

then

$$v = \frac{1}{a} \int \left[\frac{u^2 + a^2}{u^2 - 1} \right]^{\frac{1}{2}} \frac{du}{u}.$$

To carry out the integration we substitute a new letter for the radical and find

$$v = \frac{1}{2a} \log \frac{[u^2 + a^2]^{\frac{1}{2}} + [u^2 - 1]^{\frac{1}{2}}}{[u^2 + a^2]^{\frac{1}{2}} - [u^2 - 1]^{\frac{1}{2}}} - \text{arc tan} \left[\frac{1}{a}\left(\frac{u^2 + a^2}{u^2 - 1}\right)^{\frac{1}{2}}\right].$$

Replacing a by tan α and introducing φ for the arc tangent of the preceding equation, we have

$$u = \frac{\sin \alpha}{[-\cos(\alpha + \varphi)\cos(\alpha - \varphi)]^{\frac{1}{2}}}, \qquad v = \frac{\cot \alpha}{2} \log \frac{\cos(\alpha - \varphi)}{\cos(\alpha + \varphi)} - \varphi.$$

The last equations are parametric equations for the polar coördinates u, v of the projection of the helix on the xy-plane. This projection is a spiral, running from $\varphi = \frac{1}{2}\pi$, $u = 1$, $v = -\frac{1}{2}\pi$ to $\varphi = \frac{1}{2}\pi \pm \alpha$, $u = \infty$, $v = \infty$; the double sign in the last value of φ gives two spirals symmetrical with respect to the y-axis.

We find that the angle β which the tangent to the spiral projection makes with the radius vector is equal to the parameter φ, for

$$\tan \beta = u \frac{dv}{du} = \tan \varphi.$$

Since for $\varphi = \beta = \frac{1}{2}\pi$ we have $u = 1$, the spiral is tangent to this unit circle in the xy-plane, a fact which is evident geometrically.

On the hyperboloid of revolution of two sheets, meridian $u^2 - z^2/b^2 = -1$, there is no direction making with the z-axis an angle C such that $b \tan C \leqq 1$. We have only to consider the case $b \tan C > 1$ as in the preceding case. The discussion and the resulting equations are similar to those for the hyperboloid of one sheet. For the hyperboloid of two sheets

$$u = \frac{\sin \alpha}{[\cos(\alpha + \varphi)\cos(\alpha - \varphi)]^{\frac{1}{2}}}, \qquad v = \frac{\cot \alpha}{2} \log \frac{\cos(\alpha - \varphi)}{\cos(\alpha + \varphi)} - \varphi,$$

$$\varphi = \text{arc tan} \left[\frac{1}{a}\left(\frac{u^2 - a^2}{u^2 + 1}\right)^{\frac{1}{2}}\right].$$

For $\varphi = 0$, $v = 0$ and $u = a$, its smallest value; the whole extent of this spiral projection of the helix is given by φ increasing from 0 to $\frac{1}{2}\pi - \alpha$, the last-named value giving $u = \infty$, $v = \infty$. As in the previous case $\beta = \varphi$, giving, however, to this spiral a different appearance; for $u = a$, the angle $\beta = 0$, so that the spiral is tangent to the radius vector rather than perpendicular to it, as in the previous case, giving it a form resembling that of the involute of the circle—as seems geometrically plausible from the similarity in appearance of one sheet of the hyperboloid to that of the paraboloid.

4. **Surfaces determined from the projections of the helices.** When the surface of revolution is given and the helices of the surface having the same axis as the surface are required, we can apparently approach the problem only by the

method of the preceding section. We have found in several of the cases considered that the projections of the helices on a plane perpendicular to the common axis have an interest; we accordingly state our problem, so to speak, the other way round: given the projections of the helices on a surface of revolution with the same axis as the helices on a plane perpendicular to the common axis, we determine the surface. The solution of the problem so stated is very simple and in some cases advantageous as giving more easily than the method of the third section the helices on the ellipsoid and those not yet found on the hyperboloid of one sheet.

Suppose the common axis of surface and helices is the z-axis and the equations of the helix

$$x = x(t), \qquad y = y(t), \qquad z = z(t).$$

The projection on the xy-plane is given by the first two of these equations and is known; $z(t)$ is not known but

$$\frac{dz}{ds} = \frac{z'}{[x'^2 + y'^2 + z'^2]^{\frac{1}{2}}} = \cos C.$$

From this equation

$$z = \cot C \int [x'^2 + y'^2]^{\frac{1}{2}}\, dt.$$

Since $u^2 = x^2 + y^2$, we have parametric equations of the meridian of the required surface, u and z as functions of t.

We apply this method first to a case of the last section, where the projection is an involute of a circle, as an illustration and as verification of the result there found. For the projection

$$x = R(\cos t + t \sin t), \qquad y = R(\sin t - t \cos t)$$

we find $[x'^2 + y'^2]^{\frac{1}{2}} = Rt$, so that

$$z = R \cot C \frac{t^2}{2} + K, \qquad u^2 = R^2(1 + t^2)$$

and the meridian is

$$u^2 = 2R\,(z - K) \tan C + R^2 = 2pz$$

if we choose $2K = R \cot C$ and $p = R \tan C$.

Consider next the hypocycloid and the epicycloid as the projections of the helices. The standard equations of the hypocycloid, as usually given in the textbooks, are

$$x = (R - r) \cos t + r \cos \frac{R - r}{r}\, t, \qquad y = (R - r) \sin t - r \sin \frac{R - r}{r}\, t.$$

Suppose R, the radius of the "fixed" circle, not to change; these equations are those of the hypocycloid in the simplest case when $\frac{1}{2}R > r > 0$; for $R > r > \frac{1}{2}R$ they are again the equations of a hypocycloid for which the radius of the "rolling circle" is $R - r$, which is positive and less than $\frac{1}{2}R$; when $r > R$ the equations are those of an epicycloid where the radius of the rolling circle is $r - R$; finally for negative r the equations become the standard equations of the epicycloid. We may then suppose for the hypocycloid $\frac{1}{2}R > r > 0$ without other restriction than the exclusion of the trivial straight line case $R = 2r$, while for the epicycloid we replace r by $-r$.

From the equations

$$u^2 = x^2 + y^2 = (R - r)^2 + r^2 + 2r(R - r) \cos \frac{Rt}{r},$$

$$x'^2 + y'^2 = \frac{4(R - r)}{R} \sin^2 \frac{Rt}{2r}, \qquad z = \frac{4(R - r)}{R} \cot C \cos \frac{Rt}{2r}.$$

Eliminating t from the equations for u and z, we have for the meridian of the required surface of revolution

$$u^2 - \frac{R^2 \tan^2 C}{4r(R - r)} z^2 = (R - 2r)^2.$$

Without essential restriction we suppose $R - 2r = 1$ and write the equation of the meridian $u^2 - z^2/b^2 = 1$, where

$$b^2 \tan^2 C = \frac{4r(R - r)}{R^2} = 1 - \frac{1}{R} < 1.$$

The hypocycloids are then the projections of the helices of the hyperboloids of one sheet not considered in §3, the case $b \tan C < 1$. The two types of helices are separated by the case $b \tan C = 1$, true only for the rulings of the hyperboloid which do indeed satisfy the definition of a helix.

Finally, when the projections of the helices are epicycloids, the meridian of the surface is

$$u^2 + \frac{R^2 \tan^2 C}{4r(R + r)} z^2 = (R + 2r)^2,$$

obtained by changing the sign of r in the equation above. Supposing $R + 2r = 1$, we write the equation of the meridian $u^2 + z^2/b^2 = 1$, where $b^2 \tan^2 C = 4r(R + r)/R^2 = 1/R^2 - 1$. When $b = 1$, the surface is a sphere and $R = \cos C$, a result in agreement with the result of §2. For all other values of b the surface is an ellipsoid. Since a value of R, less than one, is determined for all b, C by the equation $1/R^2 - 1 = b^2 \tan^2 C$, all helices of an ellipsoid of revolution are projected on a plane perpendicular to the common axis as epicycloids.

YALE UNIVERSITY.

THE REPRESENTATION OF FUNCTIONS BY FOURIER INTEGRALS

By Alberto González Domínguez

1. **Introduction.** Let $f(t)$ be a complex-valued function of the real variable t, bounded in $(-\infty, \infty)$. Cramér [2][1] has recently established necessary and sufficient conditions in order that $f(t)$ admit almost everywhere one of the following representations:

$$f(t) = \int_{-\infty}^{\infty} e^{itx} g(x)\, dx, \tag{g}$$

$$f(t) = \int_{-\infty}^{\infty} e^{itx} dG(x). \tag{G}$$

In these formulas $g(x)$ is a complex-valued function belonging to the class $L(-\infty, \infty)$, and $G(x)$ is a complex-valued function of bounded variation in $(-\infty, \infty)$. Cramér also considered the case in which $G(x)$ is real, bounded and non-decreasing.

In this note we establish necessary and sufficient conditions for the representation almost everywhere of $f(t)$ as a Fourier integral of the above types, $G(x)$ and $g(x)$ belonging to certain other special classes. We obtain also another characteristic condition for the representation of type (g) with $g(x) \in L(-\infty, \infty)$. We also make some applications of the developed method.[2]

2. **The summation function** $s(t)$.[3] Let $s(t)$ be a function which satisfies the following conditions:

$$\int_{-\infty}^{\infty} | s(t) |\, dt < M, \tag{1}$$

$$s(t) = \int_{-\infty}^{\infty} e^{itx} K(x)\, dx, \tag{2}$$

(3) $K(x)$ is real, non-negative, even and $O(| x |^{-1-\alpha})$ as $| x | \to \infty$, α being a positive number; and

$$s(0) = \int_{-\infty}^{\infty} K(x)\, dx = 1. \tag{4}$$

Received November 13, 1939. The author is a Guggenheim Fellow.

[1] Numbers in brackets refer to the bibliography at the end of the paper.

[2] The author wishes to express his indebtedness to Professor J. D. Tamarkin, who read the manuscript and suggested improvements.

[3] These functions were already considered, before Cramér, by Bochner for a related purpose; see [1], p. 47. Cramér does not assume that $K(x)$ is even and $O(| x |^{-1-\alpha})$ as $| x | \to \infty$; but there is no practical restriction in assuming this condition, which is indeed fulfilled by all the particular functions cited by Cramér as examples (namely, the summation factors of Weierstrass, Poisson and Cesáro).

Since

$$K(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx}\, s(t)\, dt, \tag{5}$$

we see that $K(x)$ is also bounded and continuous. Furthermore, the Fourier transform $F(-y)$ of

$$e^{-itx}\, s\left(\frac{t}{n}\right)$$

is

$$F(-y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx}\, s\left(\frac{t}{n}\right) e^{ity}\, dt = nK[n(x - y)]; \tag{6}$$

and, in virtue of the hypothesis we have made on $K(x)$, it is readily seen that $nK[n(x - y)]$ is a positive kernel of a singular integral, suitable for the representation almost everywhere of any integrable function in $(-\infty, \infty)$, and of any bounded function at every point at which it has limits to the right and to the left (see [1], p. 20, and [8], p. 28).

Let us now introduce the function (considered by Cramér)

$$g(n, x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx}\, s\left(\frac{t}{n}\right) f(t)\, dt, \tag{7}$$

which is bounded and continuous.

We can now formulate our theorems.

3. **The case** $g(x) \in L(-\infty, \infty)$. As already stated, Cramér has given necessary and sufficient conditions for this type of representation. We shall give a short proof of an alternative set of conditions. In the new version, the close relationship of the theorem to certain special results of Young (see [10], p. 56) and of Fichtenholz ([4], p. 6) on Fourier series is clearly visible.

THEOREM 1. *The following conditions are necessary and sufficient for $f(t)$ to admit almost everywhere a representation of type* (g), *with $g(x) \in L(-\infty, \infty)$:*

$$\int_{-\infty}^{\infty} |\, g(n, x)\, |\, dx < M \qquad (n = 1, 2, 3, \cdots); \tag{8a}$$

(8b) $\int_S |\, g(n, x)\, |\, dx \leqq \epsilon$ *whenever the measure of the set S is less than some positive number $\delta(\epsilon)$.*

The conditions are necessary. Indeed, if we have the representation (g), with $g(x) \in L(-\infty, \infty)$, we obtain, by an application of the Parseval formula to $g(n, x)$, in view of (6)

$$g(n, x) = n \int_{-\infty}^{\infty} K[n(x - y)]\, g(y)\, dy; \tag{9}$$

also

$$(10)\qquad |\, g(n, x)\,| \leqq n \int_{-\infty}^{\infty} K[n(x-y)]\,|\, g(y)\,|\, dy,$$

$$(11)\qquad \begin{aligned} \int_{-\infty}^{\infty} |\, g(n, x)\,|\, dx &\leqq n \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} K[n(x-y)]\,|\, g(y)\,|\, dy \\ &= n \int_{-\infty}^{\infty} |\, g(y)\,|\, dy \int_{-\infty}^{\infty} K[n(x-y)]\, dx = \int_{-\infty}^{\infty} |\, g(y)\,|\, dy = M, \end{aligned}$$

so that condition (8a) is satisfied. Now $g(y) = g_1(y) + ig_2(y)$, so that (9) implies, because of the properties of the kernel $nK[n(x - y)]$ already referred to,

$$g_1(n, x) \to g_1(x), \qquad g_2(n, x) \to g_2(x)$$

almost everywhere. Now, the functions $g_1(y)$ and $g_2(y)$ being summable, we can suppose they are non-negative; we have then, in virtue of (11),

$$\int_{-\infty}^{\infty} g_1(n, x)\, dx = \int_{-\infty}^{\infty} g_1(x)\, dx,$$

and a similar equality for $g_2(x)$. But this equality, in conjunction with the previous ones, has as a consequence, in virtue of a well-known theorem of Vitali-de la Vallée Poussin (see [3], p. 13), that the integrals

$$\int_{-\infty}^{x} g_1(n, x)\, dx$$

are uniformly absolutely continuous; and so (8b) is also a necessary condition.

The conditions are sufficient. Indeed, (8a) assures us, in virtue of a theorem of Cramér (see [2], p. 202), that $f(t)$ admits almost everywhere a representation of type (G), with $G(x)$ of bounded variation in $[-\infty, \infty]$, and such that

$$G(x) = \lim_{n\to\infty} \int_{-\infty}^{x} g(n, x)\, dx = \lim_{n\to\infty} G(n, x).$$

Now, the condition (8b) tells us that the absolute continuity of the functions $G(n, x)$ is uniform, so that the limit function $G(x)$ must also be absolutely continuous, and the theorem is proved.

Remark 1. We have supposed that $f(t)$ is bounded; but the theorem is also valid if we suppose only that $f(t)s(t/n)$ is absolutely integrable for every n; and especially, as Cramér himself has remarked (see [2], pp. 199–200), if we take

$$(12)\qquad s(t) = \begin{cases} 1 - |\, t\,| & \text{for } |\, t\,| \leqq 1, \\ 0 & \text{for } |\, t\,| \geqq 1, \end{cases}$$

the theorem remains valid if we only suppose $f(t)$ integrable in every finite range. In this amplified form, Theorem 1 contains as a special case a theorem of Offord (see [7], p. 256), which asserts that if $f(t)$, integrable in every finite

range, satisfies the conditions (8a), (8b) ($s(t)$ being the function defined by (12)), then we have almost everywhere

$$f(t) = \lim_{A\to\infty} \int_{-A}^{A} \left(1 - \frac{|x|}{A}\right) e^{ixt} g(x)\, dx.$$

Offord's result follows obviously from Theorem 1 by the consistency theorem for Cesáro summability.

Remark 2. The equivalent condition found by Cramér for the case above discussed is

$$\lim_{\substack{n'\to\infty \\ n''\to\infty}} \int_{-\infty}^{\infty} | g(n', x) - g(n'', x) | \, dx = 0. \tag{13}$$

The proof of the sufficiency of (13), as given by Cramér, leaves nothing to be desired as regards simplicity. But his proof of the necessity is not quite so simple; therefore, it is perhaps not out of place to show that the present method allows us also to prove the necessity of (13) in a very simple manner.[4] We have, indeed,

$$g(n, x) - g(x) = n \int_{-\infty}^{\infty} K(nz) [g(x - z) - g(x)]\, dz,$$

$$| g(n, x) - g(x) | \leqq n \int_{-\infty}^{\infty} K(nz) \, | g(x - z) - g(x) | \, dz,$$

$$\int_{-\infty}^{\infty} | g(n, x) - g(x) | \, dx \leqq n \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} | g(x - z) - g(x) | \, dx \right] K(nz) dz. \tag{14}$$

Now the function

$$\varphi(z) = \int_{-\infty}^{\infty} | g(x - z) - g(x) | \, dx$$

is bounded, continuous, and $\varphi(0) = 0$. We see therefore (see [1], p. 20) that when n tends to infinity the last member of (14) tends to zero. Thus, $g(n, x)$ tends in mean of order one to $g(x)$, and also, as we wanted to prove,

$$\lim_{\substack{n'\to\infty \\ n''\to\infty}} \int_{-\infty}^{\infty} | g(n', x) - g(n'', x) | \, dx = 0.$$

An immediate corollary of Theorem 1 is the following

THEOREM 2. *Necessary and sufficient conditions in order that $f(t)$ admit almost everywhere a representation of type* (g), *with $g(x) \in L(-\infty, \infty)$ and essentially bounded, are the following*:

(15a) *the same as* (8a),

$$| g(n, x) | \leqq C \qquad (n = 1, 2, 3, \cdots). \tag{15b}$$

[4] See, for a particular case, Hille and Tamarkin, [5], pp. 339–340.

We have already seen that condition (15a) is necessary; and, if C is the essential upper bound of $| g(x) |$, we have in the present case

$$| g(n, x) | \leqq Cn \int_{-\infty}^{\infty} K[n(x - y)] \, dy = C.$$

The conditions are sufficient. Indeed, (15b) implies (8b) and so $f(t)$ admits, in virtue of Theorem 1, a representation of type (g), with $g(x)$ integrable in $(-\infty, \infty)$. Then we have also, as already stated,

$$\lim_{n\to\infty} g(n, x) = g(x)$$

almost everywhere; and, in virtue of (15b), the limit function $g(x)$ is also bounded almost everywhere, and the theorem is proved.

Remark. The result remains valid if $f(t)s(t/n)$ is absolutely integrable for every n; and condition (15b) alone is sufficient for $f(t)$ to admit almost everywhere the representation

$$f(t) = \lim_{A\to\infty} \int_{-A}^{A} \left(1 - \frac{|x|}{A}\right) e^{ixt} g(x) \, dx,$$

with $g(x)$ essentially bounded. This has been proved by Offord (loc. cit.) and by Verblunsky ([9], p. 19) with two special kernels; and Offord's proof can be adapted to yield the general theorem.

4. $g(x)$ **belongs to** L **and to** L^p.

THEOREM 3. *The following are necessary and sufficient conditions in order that* $f(t)$ *admit almost everywhere a representation of type* (g), *with* $g(x)$ *integrable and of p-th power integrable in* $(-\infty, \infty)$:

(16a) *the same as* (8a),

(16b) $$\int_{-\infty}^{\infty} | g(n, x) |^p \, dx < M^p \qquad (n = 1, 2, 3, \cdots).$$

We have already proved that condition (16a) is necessary. Applying now the Hölder inequality to the second member of (10) (taking into account that $K(x)$ is bounded and integrable, and so belongs to every L^p class), we get

$$| g(n, x) | \leqq \left[n \int_{-\infty}^{\infty} K[n(x - y)] \, | g(y) |^p \, dy \right]^{1/p} \left[n \int_{-\infty}^{\infty} K[n(x - y)] \, dy \right]^{(p-1)/p};$$

and it is immediately seen that the second factor is equal to 1. Raising both members to the p-th power and integrating between $(-\infty, \infty)$, we obtain, in view of the absolute integrability,

$$\int_{-\infty}^{\infty} | g(n, x) |^p \, dx \leqq \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} nK[n(x - y)] \, | g(y) |^p \, dy$$

$$= \int_{-\infty}^{\infty} | g(y) |^p \, dy \int_{-\infty}^{\infty} nK[n(x - y)] \, dx = \int_{-\infty}^{\infty} | g(y) |^p \, dy = M^p.$$

The condition (16b) is also necessary. Now as to the sufficiency, condition (16a) assures us again that $f(t)$ admits almost everywhere a representation of type (G), with $G(x)$ of bounded variation in $(-\infty, \infty)$ and such that

$$G(x) = \lim_{n\to\infty} \int_{-\infty}^{x} g(n, x)\, dx = \lim_{n\to\infty} G(n, x).$$

Now let us take a set S of non-overlapping intervals $(x_\nu, x_\nu + h_\nu)$ such that $\sum_{\nu=1}^{m} h_\nu = \delta$. We have then, in virtue of the Hölder inequality,

$$\sum_{\nu=1}^{m} | G(n, x_\nu + h_\nu) - G(n, x_\nu) | \leqq \int_S | g(n, x) | \, dx$$
$$\leqq \left[\int_S | g(n, x) |^p \, dx \right]^{1/p} \cdot \delta^{(p-1)/p} \leqq \left[\int_{-\infty}^{\infty} | g(n, x) |^p \, dx \right]^{1/p} \cdot \delta^{(p-1)/p} < M\delta^{(p-1)/p}.$$

Choosing $| S | = \delta$ sufficiently small, and making n tend to infinity, we see that $G(x)$ is absolutely continuous:

$$G(x) = \int_{-\infty}^{x} g(x)\, dx.$$

It remains to show that $g(x)$ belongs to L^p. Now, we know already that $g(n, x) \to g(x)$ almost everywhere; also $| g(n, x) |^p \to | g(x) |^p$ almost everywhere; and in virtue of Fatou's lemma

$$\int_{-\infty}^{\infty} | g(x) |^p \, dx \leqq \underline{\lim} \int_{-\infty}^{\infty} | g(n, x) |^p \, dx,$$

and the theorem is proved.

Remark. Condition (16b) alone is sufficient for $f(t)$ to admit almost everywhere the representation

$$f(t) = \lim_{A\to\infty} \int_{-A}^{A} \left(1 - \frac{| x |}{A}\right) e^{ixt} g(x)\, dx,$$

with $g(x)$ belonging to L^p. This has been proved by Offord (loc. cit.) with a special $s(t)$, and his proof can again be adapted to yield the general theorem.

5. $g(x) \in L(-\infty, \infty)$ **and of bounded variation in** $(-\infty, \infty)$. Let us now assume that the function $K(x)$ satisfies not only the conditions (2), (3) and (4), but also the following: $K(x)$ is everywhere differentiable and its derivative is everywhere bounded:

$$| K'(x) | < C.$$

Then we have the following

THEOREM 4. *The following are necessary and sufficient conditions in order that $f(t)$ admit almost everywhere a representation of type* (g), *with $g(x)$ integrable and of bounded variation in* $(-\infty, \infty)$:

(17a) *the same as* (8a),

(17b) *the same as* (8b),

$$\int_{-\infty}^{\infty} |g'(n, x)| \, dx < M \qquad (n = 1, 2, 3, \cdots). \tag{17c}$$

Proof. We have already seen that the two first conditions are necessary, and that we have

$$g(n, x) = n \int_{-\infty}^{\infty} K[n(x - y)] g(y) \, dy.$$

Now, because of the properties of $K(x)$, the differentiation under the integral sign is justified by standard theorems, and we have

$$g'(n, x) = \int_{-\infty}^{\infty} \frac{\partial}{\partial x} \{nK[n(x - y)]\} g(y) \, dy = -\int_{-\infty}^{\infty} \frac{\partial}{\partial y} \{nK[n(x - y)]\} g(y) \, dy.$$

If now we integrate the last integral by parts, we obtain (since $g(y)$ is bounded and $\lim_{|y| \to \infty} nK[n(x - y)] = 0$)

$$g'(n, x) = n \int_{-\infty}^{\infty} K[n(x - y)] \, dg(y),$$

$$|g'(n, x)| \leqq n \int_{-\infty}^{\infty} K[n(x - y)] \, |dg(y)|;$$

$$\int_{-\infty}^{\infty} |g'(n, x)| \, dx \leqq \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} nK[n(x - y)] \, |dg(y)|$$

$$= \int_{-\infty}^{\infty} |dg(y)| \int_{-\infty}^{\infty} nK[n(x - y)] \, dx = \int_{-\infty}^{\infty} |dg(y)| = C.$$

So condition (17c) is also necessary. Conversely, the two first conditions have as a consequence that $f(t)$ admits almost everywhere a representation of type (g), with $g(x)$ integrable in $[-\infty, \infty]$. It remains to show that in virtue of condition (17c), $g(x)$ is also of bounded variation in $[-\infty, \infty]$. In order to prove this, let us take an arbitrary set of non-overlapping intervals $(x_\nu, x_\nu + h_\nu)$; we have then

$$\sum_{\nu=1}^{m} |g(n, x_\nu + h_\nu) - g(n, x_\nu)| \leqq \sum_{\nu=1}^{m} \left| \int_{x_\nu}^{x_\nu + h_\nu} g'(n, x) \, dx \right|$$

$$\leqq \sum_{\nu=1}^{m} \left[\int_{x_\nu}^{x_\nu + h_\nu} |g'(n, x)| \, dx \right] \leqq \int_{-\infty}^{\infty} |g'(n, x)| \, dx < M \quad (n = 1, 2, \cdots).$$

Now, $g_n(x) \to g(x)$ almost everywhere; therefore, making $n \to \infty$ in the above inequality, we obtain

$$\sum_{\nu=1}^{m} | g(x_\nu + h_\nu) - g(x_\nu) | \leqq M.$$

Therefore, $g(x)$ is almost everywhere equal to a function of bounded variation, and the theorem is proved.

6. $G(x)$ **continuous and of bounded variation in** $[-\infty, \infty]$.

THEOREM 5. *The following are necessary and sufficient conditions for $f(t)$ to admit almost everywhere a representation of type* (G), *with $G(x)$ continuous and of bounded variation in* $[-\infty, \infty]$:

(18a) *the same as* (8a),

$$\lim_{n\to\infty} \frac{g(n, x)}{n} = 0. \tag{18b}$$

This is an immediate consequence of the theorem of Cramér that (8a) is necessary and sufficient for $f(t)$ to be a Fourier-Stieltjes integral, with $G(x)$ of bounded variation in $[-\infty, \infty]$, and of the following

THEOREM 6. *If we have*

$$f(t) = \int_{-\infty}^{\infty} e^{ixt}\, dG(x),$$

with $G(x)$ of bounded variation in $[-\infty, \infty]$, then

$$\lim_{n\to\infty} \frac{g(n, x)}{n} = \lim_{n\to\infty} \int_{-\infty}^{\infty} K[n(x - y)]\, dG(y) = K(0)[G(x + 0) - G(x - 0)].$$

In order to prove this, let us introduce the function

$$h(y) = \begin{cases} G(y) & \text{for } -\infty < y < x, \\ G(y) - G(x + 0) + G(x - 0) & \text{for } \quad x \leqq y < \infty. \end{cases}$$

We have then

$$\frac{g(n, x)}{n} = K(0)[G(x + 0) - G(x - 0)] + \int_{-\infty}^{\infty} K[n(x - y)]\, dh(y). \tag{19}$$

Let us now split the interval of integration into three parts: $(-\infty, x - \delta)$, $(x - \delta, x + \delta)$, $(x + \delta, \infty)$. The first and the third parts can then be made as small as we like for every δ by taking n sufficiently great. As regards the second part, we have

$$\left| \int_{x-\delta}^{x+\delta} K[n(x - y)]\, dh(y) \right| \leqq C \int_{x-\delta}^{x+\delta} dV(y), \tag{20}$$

where $V(y)$, the total variation of $h(y)$, is continuous at the point $y = x$. Therefore, choosing δ sufficiently small, we can make the second member of (20) arbitrarily small. Then, in view of (19), the theorem is proved.

In exactly the same manner the following theorem is proved.

THEOREM 7. *The following conditions are necessary and sufficient for $f(t)$ to admit almost everywhere a representation of type* (G), *with $G(x)$ bounded, real, non-decreasing and continuous in* $[-\infty, \infty]$:

$$g(n, x) \geqq 0 \qquad (n = 1, 2, 3, \cdots), \tag{21a}$$

$$\lim_{n\to\infty} \frac{g(n, x)}{n} = 0. \tag{21b}$$

It is sufficient to remark that, according to Cramér, condition (21a) is necessary and sufficient for $f(t)$ to admit a representation of type (G), with $G(x)$ real, bounded, and non-decreasing in $[-\infty, \infty]$, and then to apply Theorem 6.

7. An application of Cramér's function.

THEOREM 8. *If we have*

$$f(t) = \int_{-\infty}^{\infty} e^{ixt}\, dG(x),$$

with $G(x)$ (in general complex) of bounded variation in $[-\infty, \infty]$, then the following inversion formula holds:

$$\lim_{n\to\infty} \int_0^x g(n, x)\, dx = \frac{G(x + 0) + G(x - 0)}{2} - \frac{G(0+) - G(0-)}{2}. \tag{22}$$

Proof. We have

$$g(n, x) = n \int_{-\infty}^{\infty} K[n(x - y)]\, dG(y).$$

Integrating between 0 and x, and inverting the order of integration by absolute convergence, we obtain $\left(\text{introducing the function } M(x) = \int_0^x K(x)\, dx\right)$

$$\int_0^x g(n, x)\, dx = \int_{-\infty}^{\infty} M[n(x - y)]\, dG(y) - \int_{-\infty}^{\infty} M(-ny)\, dG(y).$$

If we integrate by parts the integrals of the second member, the integrated parts cancel, and we obtain

$$\int_0^x g(n, x)\, dx = n \int_{-\infty}^{\infty} K[n(x - y)]G(y)\, dy - n \int_{-\infty}^{\infty} K[-ny]G(y)\, dy.$$

Taking limits in this equality for $n \to \infty$, we get (22), in view of the properties of the kernel $nK[n(x - y)]$ already referred to.

From this inversion formula it follows immediately that two Fourier-Stieltjes

integrals of the type considered are identical if and only if the corresponding functions $G(x)$ differ by a constant. This last theorem is known (see [1], p. 68), but is usually proved only when the functions are bounded, real and non-decreasing.

Another application of the inversion formula just derived is the following theorem, which is an alternative version of the famous theorem of Paul Lévy on characteristic functions.

THEOREM 9.[5] *Let $G(x)$ be a distribution function, and let $\{G_\nu(x)\}$ be a sequence of distribution functions. Then the condition*

$$\lim_{\nu\to\infty} g_\nu(n, x) = g(n, x) \qquad (-\infty < x < \infty; n = 1, 2, 3, \cdots)$$

is necessary and sufficient in order that the sequence $\{G_\nu(x)\}$ converge essentially to $G(x)$.

The proof of this theorem is easily obtained by applying the above inversion formula and two well-known theorems of Helly, and can be omitted.

REFERENCES

1. S. BOCHNER, *Vorlesungen über Fouriersche Integrale*, Leipzig, 1932.
2. H. CRAMÉR, *On the representation of functions by certain Fourier integrals*, Trans. Amer. Math. Soc., vol. 46(1939), p. 190.
3. G. C. EVANS, *The Logarithmic Potential*, Amer. Math. Soc. Colloquium Publ., New York, 1927.
4. G. FICHTENHOLZ, *Sur l'intégrale de Poisson et quelques questions qui s'y rattachent*, Fundamenta Mathematicae, vol. 13(1929), p. 6.
5. E. HILLE AND J. D. TAMARKIN, *On the absolute integrability of Fourier transforms*, Fundamenta Mathematicae, vol. 25(1935), pp. 329–352.
6. M. JACOB, *De l'application des intégrales généralisées de Fourier au calcul des probabilités*, Comptes Rendus, vol. 188(1929), pp. 541–543 and 754–757.
7. A. C. OFFORD, *On Fourier transforms*, Trans. Amer. Math. Soc., vol. 38(1935), pp. 250–266.
8. E. C. TITCHMARSH, *Introduction to the Theory of Fourier Integrals*, Oxford, 1937.
9. S. VERBLUNSKY, *Trigonometrical integrals and harmonic functions*, Proc. Lond. Math. Soc., vol. 38(1934), pp. 1–48.
10. W. H. AND G. C. YOUNG, *On the theorem of Riesz-Fischer*, Quart. Journ. Math., vol. 44(1913), pp. 49–88.

BROWN UNIVERSITY.

[5] See [6], where a similar, but less general theorem is proved.

TWO-TO-ONE TRANSFORMATIONS

By J. H. Roberts

1. An n-to-1 transformation is one for which every inverse image consists of exactly n points. This paper is concerned with *continuous* n-to-1 transformations defined over separable metric spaces. Except in this introductory paragraph only the case $n = 2$ is considered. Now a continuous n-to-1 transformation is in some respects like an at most n-to-1 interior transformation. In particular they are alike in not altering dimension, if n is finite.[1] However, an n-to-1 continuous transformation with $n = \aleph_0$ can increase dimension at will, in contrast to the case of the at most $\aleph_0$-to-1 interior transformation. Transformations which are both n-to-1 and interior have been considered by G. T. Whyburn.[2]

Now O. G. Harrold has shown[3] that no continuous 2-to-1 transformation can be defined on an arc. The main object of the present paper is to prove the following

Principal Theorem. *There does not exist a continuous* 2-*to*-1 *transformation defined on a closed* 2-*cell.*

It seems probable that the same result holds for the closed n-cell, for every n. For this reason certain lemmas have been formulated as more general theorems, with the hope that they may be of use later on.

2. Throughout this paper we use the following notation: T is a continuous, 2-to-1 transformation defined over a space M with differing topological properties, as specified in the various theorems. In every case M is at least metric. No mention is made of the space into which M is mapped, but it can always be taken as a subset of Hilbert space. The set of inverse images is an upper semi-continuous collection G filling M, and every element of G is a pair of points. For each $x \in M$ let $s(x)$ be the point such that x, $s(x)$ is a pair of the collection. Then $T(x) = T(s(x))$. Let $f(x) = \rho(x, s(x))$, where ρ is the metric on M. The

Received November 29, 1939; presented to the American Mathematical Society, October 28, 1939.

[1] This result for the n-to-1 transformation, where n is finite, was noted by O. G. Harrold (see footnote 3), and follows from W. Hurewicz, *Über dimensionserhöhende stetige Abbildungen*, Journal für die reine und angewandte Mathematik, vol. 169(1933), pp. 71–78. An at most $\aleph_0$-to-1 interior transformation cannot increase the dimension. See P. Alexandroff, Comptes Rendus de l'Académie des Sciences de l'URSS, new series, vol. 13(1936), pp. 295–299.

[2] *Interior surface transformations*, this Journal, vol. 4(1938), p. 630.

[3] *The non-existence of a certain type of continuous transformation*, this Journal, vol. 5(1939), pp. 789–793.

transformation x into $s(x)$ takes M into itself, but is not necessarily continuous. The complicated part of the proof is concerned with showing that $s(x)$ can only have discontinuities of a certain kind.

3. LEMMA 1. *If $a \epsilon M$ and $\epsilon > 0$, then there is an open set U such that* (1) *$a \epsilon U$, and* (2) *if $x \epsilon U$ then either $|f(x) - f(a)| < \epsilon$ or else $|f(x)| < \epsilon$.*

LEMMA 2. *If $x_n \to a$ and $f(x_n) \to r$, then either $r = f(a)$ or $r = 0$.*

Both lemmas are immediate consequences of the definition of $f(x)$, in view of the fact that the collection G is upper semi-continuous.

THEOREM 1. *The functions $s(x)$ and $f(x)$ are continuous over precisely the same set.*

Proof. Now $f(x) = \rho(x, s(x))$, and since ρ is always continuous, it follows that $f(x)$ is continuous when $s(x)$ is. Conversely, suppose $s(x)$ is not continuous at $x = a$. Then there is a sequence $x_1, x_2, \cdots$ such that $x_n \to a$ but $s(x_n)$ does not approach $s(a)$. Since G is upper semi-continuous, there is a subsequence $x_1', x_2', \cdots$ such that $x_n' \to a$ and $s(x_n') \to a$. Then $f(x_n') \to 0$ as $x_n' \to a$, while $f(a) \neq 0$, so that f is not continuous at $x = a$.

THEOREM 2. *Suppose M is complete. Then the subset K of M, where $f(x)$ (and hence also $s(x)$) is continuous, is dense and open in M.*

Proof. Let U be any open set in M. The function $f(x)$ is upper semi-continuous and is bounded below (in fact $f(x) > 0$). Therefore f has a point of continuity in U, say at a. Choose $V \subset U$ so that (1) $a \epsilon V$, and (2) if $x \epsilon V$, then $|f(x) - f(a)| < \frac{1}{2}f(a)$. It will be shown that f is continuous over the open set V. Let b denote a point of V, and let a positive ϵ be given. Choose ϵ' so that $\epsilon' < \epsilon$ and $\epsilon' < \frac{1}{2}f(a)$. Let W be an open set such that (1) $b \epsilon W$ and $W \subset V$, and (2) if $x \epsilon W$, then either $|f(x) - f(b)| < \epsilon'$ or else $f(x) < \epsilon'$ (see Lemma 1). But if $x \epsilon W$, then $x \epsilon V$ and $f(x) > f(a) - \frac{1}{2}f(a) > \epsilon'$; hence for x in W, $|f(x) - f(b)| < \epsilon' < \epsilon$ and f is continuous at b.

4. THEOREM 3. *Suppose ab is an arc in $K + b$ such that $f(x) \to 0$ as $x \to b$ on ab. Then there exist two arcs axb and $a'x'b$ each lying in $K + b$, having only b in common, and such that $s(axb - b) = a'x'b - b$.*

Proof. Suppose $ab - b$ is in a component R of K. Then if $R \cdot s(R) = 0$, the result is trivial, as the given arc ab has the properties of the desired arc axb. If $R \cdot s(R) \neq 0$, we proceed as follows. First write $N = ab + s(ab - b)$. Then N has these properties:

(1) it is a continuum containing a and b; and

(2) if $x \neq b$, and $x \epsilon N$, then $s(x) \epsilon N$.

It can be shown that N contains a subcontinuum N_1 having properties (1) and (2) and being irreducible with respect to these properties. Now N_1 contains an arc axb, and hence also contains $s(axb - b) + b = a'x'b$. Since $axb + a'x'b$ has properties (1) and (2), this sum must be N_1. Suppose now axb and $a'x'b$

have a common point distinct from b. Then let y be the first such point on axb in the order a to b. It follows that y' ($y' = s(y)$) is on both axb and $a'x'b$, and we have the orders $ayy'b$ on axb and $a'y'yb$ on $a'x'b$. Let N_2 denote the sum of the subarcs ay and $y'b$ of axb and the subarcs $a'y'$ and yb of $a'x'b$. Then N_2 is a subset of N_1 having properties (1) and (2) above, and since N_1 is irreducible, it follows that $N_2 = N_1$. But this is impossible. For it requires that the subarc $y'y$ of $a'x'b$ be contained in $ay + y'b$ (on axb). But then the connected set $y'y$ is a subset of one of these arcs, and that is clearly impossible, since neither contains both y and y'. This completes the proof of Theorem 3.

5. THEOREM 4. *Let M be a complete metric space. Suppose there exists an arc pq such that*

(a) *q is not in K, but $pq - q$ is in K; and*

(b) *$f(x) \to f(q)$ as $x \to q$ on the arc pq.*

Then there exists an arc $p'q'$, and an open set V, such that

(1) *$p'q'$ is in V;*

(2) *q' is not in K, but $p'q' - q'$ is in K;*

(3) *$f(x) \to f(q') = 4\epsilon$ as $x \to q'$ on the arc $p'q'$;*

(4) *if $x \in V$, then either $f(x) < \epsilon$ or $f(x) > 3\epsilon$;*

(5) *if $p''q''$ is any arc satisfying* (1) *and* (2) *above and $f(p'') < \epsilon$, then $f(x) \to 0$ as $x \to q''$ on the arc $p''q''$.*

Proof. Suppose the theorem is false. Let p_1q_1 be an arc such that (1) q_1 is not in K but $p_1q_1 - q_1$ is in K, and (2) $f(x) \to f(q_1)$ as $x \to q_1$ on the arc p_1q_1. Let $4\epsilon_1 = f(q_1)$ and let U denote any open set containing q_1. Next let U_1 be an open set in M such that $\overline{U}_1 \subset U$, diameter of $U_1 < \epsilon_1$, $q_1 \in U_1$, and if $x \in U_1$ then $f(x) < \epsilon_1$ or else $f(x) > 3\epsilon_1$, but in either case $f(x) < 5\epsilon_1$. Let p_1' be a point on the arc p_1q_1 so that $p_1'q_1$ is a subarc of p_1q_1 lying in U_1. Then there is an arc p_2q_2 in U_1 such that p_2q_2 satisfies (1) and (2), $f(p_2) < \epsilon_1$ but $f(x)$ does not approach zero as $x \to q_2$ on the arc p_2q_2. Then $f(x) \to f(q_2)$ as $x \to q_2$ on this arc. Consequently $f(q_2) < \epsilon_1$.

Now let $4\epsilon_2 = f(q_2)$, and let U_2 be an open set such that $\overline{U}_2 \subset U_1$, diameter of $U_2 < \epsilon_2$, $q_2 \in U_2$, and if $x \in U_2$ then $f(x) < \epsilon_2$ or else $f(x) > 3\epsilon_2$, but in either case $f(x) < 5\epsilon_2$. Continuing in this way we obtain an infinite sequence of open sets $U_1, U_2, \cdots$, and positive numbers $\epsilon_1, \epsilon_2, \cdots$, such that for every k (a) $U_k \supset \overline{U}_{k+1}$, (b) diameter of $U_k < \epsilon_k$, (c) $4\epsilon_{k+1} < \epsilon_k$, and (d) for $x \in U_k$, $f(x) < 5\epsilon_k$. Let x' be a point common to $U_1, U_2, \cdots$. Then $f(x') < 5\epsilon_k$ for every k, hence $f(x') = 0$. This is a contradiction, and therefore the theorem has been proved.

6. THEOREM 5. *Suppose M is a closed 2-cell, R is a component of K, $R \cdot s(R) = 0$, and q is a boundary point of R which is accessible from R by an arc pq along which $f(x) \to 0$ as $x \to q$. Suppose furthermore that there is an open set W containing q such that if uv is any arc having these properties:*

(a) *$u \in R \cdot W$ and, on pq, v is on the boundary of R; and*

(b) *uv is in W, and $uv - v$ is in R;*

then $f(x) \to 0$ as $x \to v$ on the arc uv. Then if q is not on the boundary of M, there exists, for every positive ϵ, a simple closed curve J such that

(1) *q is in the interior of J;*

(2) *the diameter of J plus its interior $< \epsilon$;*

(3) *J is the sum of two arcs axb and ayb, where a and b are in R and $s(R)$, respectively, and $s(axb - a - b) = ayb - a - b$.*

If q is on the boundary of M, then there is an arc J which separates M such that

(1) *the component of $M - J$ containing q is of diameter less than ϵ;*

(2) *J is the sum of 2 arcs xa and ya, where a is on the boundary of R, $xa - a$ and $ya - a$ are in R and $s(R)$, respectively, and $s(xa - a) = ya - a$.*

Proof. A detailed proof will be given only for the case where q is in the interior of M. Suppose pq is an arc lying in $R + q$ and such that $f(x) \to 0$ as $x \to q$ on this arc. Then $s(pq - q) + q$ is an arc $p'q$ lying in $s(R) + q$. Clearly there exist arcs rp and $p't$ such that $rp + pq + qp' + p't$ is an arc rt having its end-points, and only these points, on the boundary of M. Let U_1 and U_2 denote the 2 components of $M - rt$. For each positive integer n let $x_n y_n z_n$ be an arc in $U_1 + x_n + z_n$ whose diameter $< 1/n$, and such that x_n is on $pq - q$ and $z_n = s(x_n)$. Let a_n be the first point, in the order $x_n z_n$, of the boundary of R on this arc. Then for n sufficiently large the arc $s(x_n a_n - a_n) + a_n = z_n a_n$ lies in $U_1 + z_n$. Then $x_n a_n z_n$ is an arc in $U_1 + a_n + z_n$ such that $s(x_n a_n - a_n) = z_n a_n - a_n$. Similarly, in $U_2 + x_n + z_n$ there is an arc $x_n b_n z_n$ with similar properties. The sum of the arcs $x_n a_n z_n$ and $x_n b_n z_n$ is the desired curve J, where a_n and b_n are the points a and b.

7. We come now to the proof of the principal theorem (see §1). The proof is indirect. In the remainder of the paper it is assumed that M is a closed 2-cell, T is a continuous 2-to-1 transformation defined on M, and f, s, and K are as defined in §§2, 3.

8. *Let R be a component of K and suppose that at least three boundary points of R are accessible from R by arcs along which $f(x) \to 0$ as x approaches the boundary of R. Then $R \cdot s(R) = 0$.*

We suppose the assertion is false. Then since s is continuous over R, the set $R + s(R)$ is connected, hence is R. Since in addition s has period 2, it follows that $s(R) = R$.

There is a simple closed curve J ($J = prp'r'p$) in R, where $s(prp') = p'r'p$, and such that no point of J is on the boundary of M. For let aa' be an arc in R joining a and $s(a)$ and containing no boundary point of M. Let x be a variable point on aa', and consider the two arcs ax and $s(ax)$. For x near enough to a these have no common point, but for $x = a'$ they do have a common point. Let p' be the first point on aa' such that ap' and $s(ap')$ intersect. Then p' and p ($p = s(p')$) are the two common points, and if r is on aa' between p and p', the set $prp' + s(prp')$ is the desired curve J.

Now $M - J = C_1 + C_2$, where C_1 and C_2 are connected. One of C_1 and C_2 contains at least two points on the boundary of R which are arc-wise accessible from R by arcs along which $f(x) \to 0$ as x approaches the boundary. Suppose C_1 has this property, and let e and g be two such points in C_1. Let $D_i = R \cdot C_i$ $(i = 1, 2)$. Now D_1 and D_2 are connected. For let xy denote an arc in R connecting any two points x and y of D_i. Then if xy is not in D_i, it must intersect J. But since J is in the open set R, the arc xy can be modified to run in R near J without crossing it, giving rise to a new arc connecting x and y in D_i.

Let ae and $a'e$ be arcs (see Theorem 3) such that (a) $s(ae - e) = a'e - e$, (b) $ae \cdot a'e = e$, (c) a and a' are on J and are the only points of these arcs on J, and (d) $f(x) \to 0$ as $x \to e$ along ae. Then as $x \to e$ along the arc ae, we have $s(x) \to e$ along $a'e$, whence $s(x)$ cannot be in D_2, and thus $s(D_1) = D_1$. Suppose (case 1) that the arc $ae + a'e$ does not intersect the boundary of M. Then it separates C_1 into two components C_3 and C_4. Let $D_i = R \cdot C_i$ $(i = 3, 4)$, and as above the set D_i is connected. Then $s(D_3) = D_4$, since s takes the boundary of C_3 into the boundary of C_4 (except for the point e). But this is impossible. For the point g is in C_3 or C_4. (For definiteness suppose it is in C_3.) Then for some method of approach we will have $x_n \in D_3$, $x_n \to g, f(x_n) \to 0$, whence $s(x_n) \to g$ and therefore $s(x_n)$ is in D_3 for some n. But then $s(D_3) = D_3$, and we have a contradiction.

We now suppose (case 2) that the arc $ae + a'e$ intersects the boundary of M. Of the 2 arcs on J with end-points a and a' one of them, call it aua', together with $ae + a'e$ makes a simple closed curve J' with $J - aua'$ in its exterior. Let C_3 be the interior of J' and let $C_4 = C_1 - C_1 \cdot \bar{C}_3$. Let $D_3 = R \cdot C_3$. As before, D_3 is connected. Let $D_4 = R \cdot C_4$. We cannot at this stage prove that D_4 is connected. Now D_3 is bounded by $aua' + ae + a'e$; therefore $s(D_3)$ is bounded by $a'u'a + ae + a'e$. Then $s(D_3)$ is in C_1 and $D_3 \cdot s(D_3) = 0$. But then $s(D_3) = D_4$ and $s(D_4) = D_3$. But this is impossible as in case 1 (on account of the point g). This completes the proof.

9. *There does not exist an arc pq such that*
(1) *$pq - q$ is in K but q is not in K; and*
(2) *$f(x) \to f(q)$ as $x \to q$ on the arc pq.*

We suppose the contrary. Then by Theorem 4 there exists an open set V and an arc pq such that properties (1)–(5) of Theorem 4 hold, where pq takes the place of $p'q'$ and $f(q) = 4\epsilon$. Since f is not continuous at q, there is a point c in K and an arc cq lying in V such that $f(c) < \epsilon$. Let R be the component of K containing c. We next prove that q is on the boundary of R and is accessible from R by an arc along which $f(x) \to 0$ as $x \to q$.

We note first that it follows readily from the assertion in §8 that $R \cdot s(R) = 0$. Now let E denote the set of all x in $R + s(R)$ for which $f(x) < \epsilon$. Let t be the last point of E on the arc cq. Suppose that t is not accessible from R by an arc ct such that $f(x) \to 0$ as $x \to t$ on ct. Then t is not accessible by any arc dt lying in V and in $R + s(R) + t$ and containing a point of E. It follows that there is

an infinite sequence $c_1, c_2, c_3, \cdots$ such that[4] (1) $c_n \in R \cdot V$ and $f(c_n) < \epsilon$, (2) $c_n \to t$ as $n \to \infty$, and (3) there is a fixed positive δ such that every arc joining c_i and c_j $(i \neq j)$ in R has diameter greater than δ.

There exist[5] 3 circles P_1, P_2, and P_3 with center t and interiors W_1, W_2, and W_3, and an integer N such that

(1) $V \supset \overline{W}_1$ and $W_i \supset \overline{W}_{i+1}$ $(i = 1, 2)$;

(2) if $n > N$ there is an arc $c_n d_n e_n$ in R and in $\overline{W}_1$, where e_n and d_n are on the boundaries of W_1 and W_2, respectively, and c_n is in W_3, and

(3) if $n > N$, $m > N$, and $n \neq m$, then no component of $R \cdot W_1$ contains both c_n and c_m.

Suppose $n > N$. Let x_n and y_n be the first points of the boundary of R on the circle P_2 starting from d_n in the two senses, and let $x_n d_n y_n$ denote the indicated arc on P_2. Then $s(x_n d_n y_n - x_n - y_n) + x_n + y_n$ is an arc $x_n d_n' y_n$ in $s(R) + x_n + y_n$. Since it does not intersect $c_n d_n e_n$, there is a positive δ_1 independent of n such that the diameter of $x_n d_n' y_n > \delta_1$. But the diameter of $x_n d_n y_n \to 0$ as $n \to \infty$. If we drop to a subsequence, we may suppose $\overline{\lim\limits_{n\to\infty}}\, x_n d_n y_n$ is a point r. Then $\overline{\lim\limits_{n\to\infty}}\, x_n d_n' y_n$ is contained in $r + s(r)$. But this is clearly impossible, and thus t is accessible by an arc ct lying in $R + t$ and such that $f(x) \to 0$ as $x \to t$ on ct.

Suppose now $t \neq q$. Suppose t is not on the boundary of M. Then by Theorem 5 there is a simple closed curve J, such that (1) t is in the interior of J, (2) q is in the exterior of J, and (3) $J = axb + ayb$, where axb is in $R + a + b$, ayb is in $s(R) + a + b$, $f(z) \to 0$ as $z \to a$ or as $z \to b$ on axb or on ayb. Then clearly some point of the arc tq is on J, and such a point is in $\bar{E}$, contrary to the supposition that t is the last point of $\bar{E}$ on cq. If t is on the boundary of M, we get a contradiction in a similar way, again using Theorem 5. Thus $t = q$. Then let J be a curve as above, inclosing q but leaving p outside. Then there is a point z of K on $J \cdot pq$, for which $f(z) > 3\epsilon$. But as $w \to a$ on J we have $f(w) \to 0$, and $f(w) < \epsilon$ or $f(w) > 3\epsilon$. This contradiction completes the proof of the assertion in this section.

10. *There are at least* 3 *points in* $M - K$.

First of all $K \neq M$, for if so, then the transformation x into $s(x)$ defines a continuous mapping of the closed 2-cell M into itself without fixed point. If a is not in K, then $s(a)$ is not in K (see Theorem 2), whence $M - K$ contains at least two points, a and $a' = s(a)$. Suppose $M - K = a + a'$. Then $\lim\limits_{x\to a} f(x) = 0$ and $\lim\limits_{x \to a'} f(x) = 0$. It follows (see Theorem 3) that there is a simple closed curve J, which is the sum of two arcs $axya'$ and $ax'y'a'$, where

[4] Since f is not continuous at t, there is such a sequence of c_n's in E for which $f(c_n) \to 0$. Either infinitely many of these are in R, or in $s(R)$, and we assume the first alternative. The other case is treated in exactly the same way, since $s(s(R)) = R$.

[5] This argument is similar to that by which one proves that a boundary point of a complementary domain of a plane continuous curve is arc-wise accessible from that domain.

$s(axya' - a - a') = ax'y'a' - a - a'$. Let I be the interior of J. Then since $s(J) = J$, it follows that $s(I) = I$ or else $s(I) = M - \bar{I}$. But if $s(I) = M - \bar{I}$, then s is a topological mapping over I, and this is impossible since $M - \bar{I}$ is not an open 2-cell. Hence $s(I) = I$. Then (compare the description of J in §8 above) there is a simple closed curve L with interior Q such that $L + Q$ is in I and $s(L) = L$. Then as before $s(Q) = Q$. But then the transformation x into $s(x)$ is a continuous mapping of the closed 2-cell $L + Q$ into itself without fixed point, and we have reached a contradiction.

11. In this section we complete the proof of the principal theorem.

Let R be a component of K. By §10 the boundary of R contains at least 3 points. By §9 $f(x) \to 0$ as x approaches the boundary of R on an arc. Hence by §8 $R \cdot s(R) = 0$. By an argument like that in §9 we can show that every boundary point of R is accessible from R. Suppose there exists a point p of M not in $\bar{R} + s(R)$, and let pq be an arc in M having only q in $\bar{R} + s(R)$. Then q is accessible from R by an arc cq. But we get a contradiction by applying Theorem 5. Thus $\bar{R} + s(R) = M$.

Let H denote the boundary of R. Suppose $H = H_1 + H_2$, mutually separated sets. Let a and b denote points of H_1, H_2, respectively, and let axb and ayb be arcs in $R + a + b$ and $s(R) + a + b$, respectively. Let I denote the interior of the simple closed curve $axb + ayb$. Neither of the closed sets $H_1 \cdot \bar{I}$ and $H_2 \cdot \bar{I}$ separates x from y in $\bar{I}$. Hence their sum does not, and there is an arc xy in $\bar{I}$ not intersecting H. But this is impossible, since x and y are in R and $s(R)$, respectively. Therefore H is connected.

Now if every pair a, b in H separates H, then H is a simple closed curve. Suppose some pair a, b does not, and let axb and ayb be arcs as above. Then if d is any point of $H - a - b$, there exist arcs xd in $R + d$ and yd in $s(R) + d$, where

$$(xd + yd) \cdot (axb + ayb) = x + y.$$

It follows that d separates H, whence H is an arc from a to b.

Now $s(H) = H$. Hence H cannot be an arc (see §1), for the original transformation T is exactly 2-to-1 and continuous over H. Thus H is a simple closed curve. Now s is a topological transformation of R into $s(R)$. One of R, $s(R)$ is the interior of H, the other is the exterior of H with respect to M. But these sets are not homeomorphic. This final contradiction completes the proof of the principal theorem.

DUKE UNIVERSITY.

THE FIRST VARIATION IN MINIMAL SURFACE THEORY

By Marston Morse

Introduction. We are concerned with harmonic surfaces S bounded by $n + 1$ non-intersecting simple closed curves[1] γ_k in a Euclidean m-space (x). The coördinates x^j on such surfaces, which are necessarily orientable and of genus 0, shall be harmonic functions of parameters (u, v) ranging over a connected region[2] B in the (u, v)-plane bounded by $n + 1$ circles $C_0, \cdots, C_n$. Let (r_k, θ_k) be polar coördinates in the (u, v)-plane with pole at the center (u_k, v_k) of C_k. Let $g_k(\theta_k)$ be a vector[3] defined on C_k giving an admissible representation of γ_k $(k = 0, 1, \cdots, n)$. Let $h(u, v)$ be the vector defining S. We suppose that $h(u, v)$ is continuous on the closure $\bar{B}$ of B and that its boundary values, represented in terms of the respective angular coördinates θ_k, determine admissible representations of the curves (γ) of the form

$$[g_0(\theta_0), \cdots, g_n(\theta_n)] = (g). \tag{0.1}$$

Conversely, each admissible representation (0.1) of the set (γ) determines a harmonic surface S defined over B. We admit no other surfaces.

Let $h(u, v)$ represent the harmonic surface defined by an admissible representation (g) of the curves (γ) and let $D(g)$ be one-half the sum (finite or infinite) of the classical[4,5] Dirichlet integrals of the components $h^i(u, v)$ of $h(u, v)$. We term $D(g)$ the *Dirichlet sum*. The radii σ_k of the circles C_k and the coördinates (u_k, v_k) of their centers will be called the *circle parameters* of B. Set[6]

$$(\sigma_1, u_1, v_1, \cdots, \sigma_n, u_n, v_n) = (\eta).$$

Received April 10, 1940.

[1] We start with a 1-1 representation of γ_k on a circle C. We shall admit any other representation of γ_k which is obtained from the given representation by a monotone transformation (not necessarily 1-1) of C into itself.

[2] Although we restrict ourselves to surfaces of the topological type of B, the methods introduced are of such general character as to admit extension to the most general type of representation. For such extensions it is sufficient that the boundaries γ_k of the type regions B be circles. It is even sufficient that there exist a conformal map of a neighborhood of γ_k on B into an annular region, γ_k going into a circle under this conformal map. The introduction of subregions of multiply-sheeted Riemann surfaces naturally makes no difficulty.

[3] Curves and surfaces in our Euclidean m-space will invariably be represented as vectors as will the various Fourier coefficients which will presently enter.

[4] Jesse Douglas I: *Solution of the problem of Plateau,* Transactions of the American Mathematical Society, vol. 33(1931), pp. 263–321. Douglas II: *The problem of Plateau for two contours,* Journal of Mathematics and Physics, vol. 10(1931), pp. 315–359. Douglas III: *Minimal surfaces of higher topological structure,* Transactions of the American Mathematical Society, vol. 39(1939), pp. 205–298. Various additional references are here given.

[5] T. Radó, *On the problem of Plateau,* Ergebnisse der Mathematik, vol. 2(1933).

[6] We do not include the parameters (σ_0, u_0, v_0) in the set (η) because without changing the value of $D(g)$, C_0 can be taken as the unit circle with center at the origin.

We shall regard D not only as a function of (g), but also as a function of (η). The representation (g) will be a representation on the respective circles defined by (η). The fundamental theorem is as follows.

A necessary and sufficient condition that a harmonic surface, for which D is finite, be minimal is that the first variation of D be identically zero.

We understand thereby that the first variation vanishes for a class of variations now to be defined.

Let (g) be an admissible representation of the circles (γ) in terms of the respective angles $(\theta_0, \cdots, \theta_n) = (\theta)$. We shall admit transformations of θ_k of the form

$$\mu_k = \theta_k + e_k\lambda_k(\theta_k), \tag{0.2}$$

where $\lambda_k(\theta_k)$ is analytic and has a period 2π, and e_k is a constant restricted by a condition of the form

$$| e_k | \leqq e_k(\lambda_k) < 1 \tag{0.3}$$

to be so small that for each value of θ_k

$$1 + e_k\lambda_k'(\theta_k) \neq 0. \tag{0.4}$$

Here $e_k(\lambda_k)$ is a constant dependent on λ_k. Subject to (0.2) we obtain a new[7] representation (p) of (γ) such that

$$p_k(\mu_k, e_k) \equiv g_k(\theta_k). \tag{0.5}$$

The boundary vectors $[p(\theta, e)]$ define a new harmonic surface $H(u, v, e)$. The corresponding Dirichlet sum $D(p)$ will be written as a function

$$D(g, \eta, \lambda, e), \qquad | e_k | \leqq e_k(\lambda_k), \tag{0.6}$$

where

$$(e_0, \cdots, e_n) = (e), \qquad (\lambda_0, \cdots, \lambda_n) = (\lambda).$$

The differential

$$V = \sum_i D_{\eta_i} d\eta_i + \sum_k D_{e_k} de_k \qquad (i = 1, \cdots, 3n; k = 0, \cdots, n), \tag{0.7}$$

evaluated[8] at $e = 0$, will be called the *first variation* of D at (g, η), and the coefficients in this differential will be called the *variational coefficients*. We shall say that the first variation of D at (g, η) is identically zero if these coefficients vanish for all admissible sets (λ) and for the given set (g, η). It is in this sense that the fundamental theorem is to be understood.

This theorem has been completely proved only for the case of one boundary

[7] We term (p) the (λ, e) "transform" of (g).

[8] Subscripts used to indicate partial differentiation shall always refer to the *original* arguments of the function. Thus if $f(x, y)$ is given, $f_x(x, x)$ means the x-derivative of $f(x, y)$ evaluated for $y = x$.

curve. The only proof known to the author is that[9] in Douglas I. When the class of admissible surfaces is enlarged to include surfaces which are no longer harmonic, Radó (loc. cit.), Courant[10] and others[11] have given proofs of the corresponding theorem. But the theorem here stated does not follow in any immediate way from the proofs by Radó and Courant. If one adds the hypothesis that the given surface is a minimizing surface, then connections can be easily made between the two types of proof, but the requirements of the theory in the large preclude such a hypothesis. This is not to say that it is impossible to proceed in the general theory except in the class of harmonic surfaces, but rather to say that in the present state of development, the class of harmonic surfaces forms the simplest and most elegant medium for the general theory.

Our proof of the theorem may be briefly described as follows. The harmonic vector $h(u, v)$ is the real part of a vector $F(w)$, whose components are analytic functions of the complex variable $w = u + iv$, the imaginary part in general being multiple-valued. With Weierstrass we introduce the scalar product[12]

$$f(w) = F'(w) \cdot F'(w), \tag{0.8}$$

whose vanishing is the condition that $h(u, v)$ be minimal. In case the boundary representation (g) is analytic, classical methods yield formulas for the variational coefficients involving $f(w)$ and (g, η, λ), and the theorem follows with ease. Our contribution is to show that when $D(g)$ is finite the variational coefficients are *continuous* in (g, η, λ) provided the metric for the sets (g, η, λ) is suitably defined. An obvious limiting process then enables us to infer the theorem in the general case as a consequence of the formulas in the special analytic case.

1. **Continuity of the variational coefficient when** $n = 0$. When $n = 0$, there is just one boundary curve $\gamma_0 = \gamma$ represented on the circle $C_0 = C$. Without loss of generality we can suppose that C is the unit circle with center at the origin. For the Dirichlet sum $D(g)$ is independent of the position of the circle C since any circle is obtainable from any other by conformal transformations of the plane. The circle parameters (η) do not enter in the case $n = 0$.

Let (r, θ) be polar coördinates with pole at the origin, and let $g(\theta)$ be an admis-

[9] In the case of two contours (Douglas II) there is a serious gap on page 328 in inferring the fundamental formula (2.18) from the preceding formula. A similar difficulty arises on page 248 of Douglas III. That this gap can be bridged by an extended use of the appropriate theta functions seems most likely. Douglas is of like opinion. Our proofs make no use of theta functions.

[10] R. Courant, *Plateau's problem and Dirichlet's principle*, Annals of Mathematics, vol. 38(1937), pp. 679–724.

[11] M. Shiffman, *The Plateau problem for non-relative minima*, Annals of Mathematics, vol. 40(1939), pp. 834–854.

[12] We term $f(w)$ the *Weierstrass modul* belonging to the region B and the boundary values (g).

sible representation of γ as an image of C. As shown in Douglas I the Dirichlet sum $D(g)$ has the form

$$(1.0)\qquad A(g) = \frac{1}{16\pi}\iint_T \frac{[g(\alpha) - g(\beta)]^2}{\sin^2[\frac{1}{2}(\alpha - \beta)]}\,d\alpha\,d\beta,$$

where T is the domain[13]

$$(1.1)\qquad \begin{gathered} 0 \leqq \alpha \leqq 2\pi, \\ \beta - \pi \leqq \alpha \leqq \beta + \pi. \end{gathered}$$

The integral $A(g)$ has a singularity on T when $\alpha = \beta$. Let T_m be the subdomain of T on which

$$|\alpha - \beta| > \frac{1}{m},$$

and let $A^{(m)}(g)$ be the value of the integral (1.0) when T is replaced by T_m. By definition

$$A(g) = \lim_{m=\infty} A^{(m)}(g).$$

In the case $n = 0$ there is but one variational coefficient, namely D_e. We shall obtain a formula for this coefficient. Let $\lambda(\theta)$ be real and analytic in θ with a period 2π. Set

$$(1.2)\qquad \mu = \theta + e\lambda(\theta), \qquad |e| \leqq e(\lambda),$$

where $e(\lambda)$ is a positive constant so small that the θ-derivative of the right member of (1.2) never vanishes. Under (1.2) let

$$p(\mu, e) \equiv g(\theta),$$

and let $H(u, v, e)$ be the harmonic surface determined by (p) for each fixed e. We have

$$H(\cos\mu, \sin\mu, e) \equiv p(\mu, e).$$

The corresponding integral $A^{(m)}(p)$ has the form

$$(1.3)\qquad A^{(m)}(p) = \frac{1}{16\pi}\iint_{T_m} \frac{[p(\mu, e) - p(\nu, e)]^2}{\sin^2[\frac{1}{2}(\mu - \nu)]}\,d\mu\,d\nu.$$

Upon setting

$$(1.4)\qquad \mu = \alpha + e\lambda(\alpha), \qquad \nu = \beta + e\lambda(\beta),$$

[13] Douglas takes T as the domain $0 \leqq \alpha \leqq 2\pi$, $0 \leqq \beta \leqq 2\pi$. This yields the same value of $D(g)$ because of the periodicity of the integrand. From the point of view of a treatment of the singularities of the integrand our choice of T seems simpler.

(1.3) takes the form

$$A^{(m)}(p) = \frac{1}{16\pi} \iint_{T_m} \frac{[g(\alpha) - g(\beta)]^2[1 + e\lambda'(\alpha)][1 + e\lambda'(\beta)]}{\sin^2 [\frac{1}{2}(\alpha - \beta) + \frac{1}{2}e(\lambda(\alpha) - \lambda(\beta))]} \, d\alpha \, d\beta. \tag{1.5}$$

We shall assume that $A(g)$ is finite and obtain a formula for $A(p) - A(g)$. To that end we shall analyze the denominator of the integrand in (1.5).

Set

$$\omega = \tfrac{1}{2}[\lambda(\alpha) - \lambda(\beta)],$$

and note that on T, for $\alpha \neq \beta$ and $e \neq 0$,

$$\begin{aligned} &\sin [\tfrac{1}{2}(\alpha - \beta) + e\omega] \\ &= \sin \tfrac{1}{2}(\alpha - \beta) \left[\cos e\omega + e \cos \tfrac{1}{2}(\alpha - \beta) \frac{\sin e\omega}{e \sin \frac{1}{2}(\alpha - \beta)}\right]. \end{aligned} \tag{1.6}$$

Observe that

$$\frac{\sin e\omega}{e \sin \frac{1}{2}(\alpha - \beta)} = \frac{\sin e\omega}{e\omega} \frac{\omega}{\alpha - \beta} \frac{\alpha - \beta}{\sin \frac{1}{2}(\alpha - \beta)}. \tag{1.7}$$

Recall that $\lambda(\alpha)$ is analytic. The form of the right member of (1.7) shows that the function thereby defined is analytic in (α, β, e) for (α, β) on T, except for removable singularities when $\alpha = \beta$ or $e = 0$. Hence (1.6) yields the relation

$$\sin [\tfrac{1}{2}(\alpha - \beta) + e\omega] = \sin \tfrac{1}{2}(\alpha - \beta) [1 + ek(\alpha, \beta, e)], \tag{1.8}$$

where $k(\alpha, \beta, e)$ is analytic in its arguments. Upon differentiating the respective members of (1.8) with respect to e and setting $e = 0$, we find that

$$\omega \cos \tfrac{1}{2}(\alpha - \beta) = k(\alpha, \beta, 0) \sin \tfrac{1}{2}(\alpha - \beta), \tag{1.9}$$

a fact of use later.

To continue, we need to know that on T,

$$1 + ek(\alpha, \beta, e) \neq 0, \qquad |e| \leqq e(\lambda). \tag{1.10}$$

We can infer this from (1.8). For

$$\sin [\tfrac{1}{2}(\alpha - \beta) + e\omega] = 0$$

only if

$$\tfrac{1}{2}(\alpha - \beta) + e\omega \equiv 0 \qquad (\text{mod } \pi),$$

or only if

$$\alpha + e\lambda(\alpha) \equiv \beta + e\lambda(\beta) \qquad (\text{mod } 2\pi). \tag{1.11}$$

Since $\alpha + e\lambda(\alpha)$ is an increasing function of α for $|e| \leqq e(\lambda)$, (1.11) holds only if $\alpha \equiv \beta$ (mod 2π). But (1.10) holds even when $\alpha \equiv \beta$, as one infers from the first order vanishing of the left member of (1.8) when $\alpha \equiv \beta$.

By virtue of the binomial theorem we can write

$$
(1.12)\qquad \begin{aligned}[1 + ek(\alpha, \beta, e)]^{-2} &= 1 - 2ek(\alpha, \beta, e) + 3e^2k^2(\alpha, \beta, e) + \cdots \qquad (|e| \leqq e(\lambda))\\ &= 1 - 2ek(\alpha, \beta, 0) + e^2B(\alpha, \beta, e),\end{aligned}
$$

where $B(\alpha, \beta, e)$ is analytic in its arguments. For brevity we set

$$
\frac{[g(\alpha) - g(\beta)]^2}{\sin^2 [\frac{1}{2}(\alpha - \beta)]} = K(\alpha, \beta).
$$

Referring to (1.5), (1.8) and (1.12), we see that

$$
(1.13)\qquad \begin{aligned} A^{(m)}(p) &= \frac{1}{16\pi} \iint_{T_m} K(\alpha, \beta)[1 - 2ek(\alpha, \beta, 0) + e^2B(\alpha, \beta, e)] \\ &\qquad\qquad \cdot [1 + e\lambda'(\alpha)][1 + e\lambda'(\beta)]\, d\alpha\, d\beta \\ &= A^{(m)}(g) + eS^{(m)}(g, \lambda) + e^2r^{(m)}(g, \lambda, e),\end{aligned}
$$

where

$$
(1.14)\qquad S^{(m)}(g, \lambda) = \frac{1}{16\pi} \iint_{T_m} K(\alpha, \beta)[\lambda'(\alpha) + \lambda'(\beta) - 2k(\alpha, \beta, 0)]\, d\alpha\, d\beta,
$$

$$
(1.15)\qquad r^{(m)}(g, \lambda, e) = \frac{1}{16\pi} \iint_{T_m} K(\alpha, \beta) W(\alpha, \beta, e)\, d\alpha\, d\beta,
$$

and where $W(\alpha, \beta, e)$ is analytic for $|e| \leqq e(\lambda)$.

The integral

$$
(1.16)\qquad A^{(m)}(g) = \frac{1}{16\pi} \iint_{T_m} K(\alpha, \beta)\, d\alpha\, d\beta
$$

converges by hypothesis as m becomes infinite. The integrals $S^{(m)}(g, \lambda)$ and $r^{(m)}(g, \lambda, e)$ also converge, as one sees by comparison with the integral $A^{(m)}(g)$, noting that the factors

$$
|\lambda'(\alpha) + \lambda'(\beta) - 2k(\alpha, \beta, 0)|, \qquad |W(\alpha, \beta, e)|
$$

are bounded. Upon letting m become infinite in (1.13), we infer that $A(p)$ is finite and that

$$
(1.17)\qquad A(p) = A(g) + eS(g, \lambda) + e^2r(g, \lambda, e),
$$

where $S(g, \lambda)$ and $r(g, \lambda, e)$ are of the form (1.14) and (1.15) respectively with T_m replaced by T.

We note that

$$
r(g, \lambda, e) \leqq LA(g),
$$

where L is a bound for $|W(a, b, e)|$. Recalling that $A(g)$ is finite, we see that

$$A_e(p) = S(g, \lambda) \qquad (e = 0).$$

Upon referring to (1.9) and (1.14), we obtain the following theorem.

THEOREM 1.1. *When $A(g)$ is finite, the variational coefficient $A_e(p)$ has the form*

$$S(g, \lambda) = \frac{1}{16\pi} \iint_T \frac{[g(\alpha) - g(\beta)]^2}{\sin^2 [\frac{1}{2}(\alpha - \beta)]} E(\alpha, \beta)\, d\alpha\, d\beta,$$

where

$$E(\alpha, \beta) = \lambda'(\alpha) + \lambda'(\beta) - [\lambda(\alpha) - \lambda(\beta)] \cot [\tfrac{1}{2}(\alpha - \beta)] \qquad (\alpha \not\equiv \beta, \bmod 2\pi).$$

We shall show that the integrand of $S(g, \lambda)$ has a removable singularity when $\alpha = \beta$. On T set

$$A(\alpha, \beta) = \tfrac{1}{2}(\alpha - \beta) \cot [\tfrac{1}{2}(\alpha - \beta)] \qquad (\alpha \neq \beta) \tag{1.18}$$

with $A(\alpha, \alpha) \equiv 1$. So defined $A(\alpha, \beta)$ is analytic. Moreover, $A(\alpha, \beta)$ is an even function of $(\alpha - \beta)$ so that $A_\beta(\alpha, \alpha) \equiv 0$.

We have

$$\lambda(\alpha) - \lambda(\beta) = (\alpha - \beta) \int_0^1 \lambda'[\alpha + t(\beta - \alpha)]\, dt,$$

and making use of (1.9) and (1.18) we see that

$$E(\alpha, \beta) = \lambda'(\alpha) + \lambda'(\beta) - 2A(\alpha, \beta) \int_0^1 \lambda'[\alpha + t(\beta - \alpha)]\, dt. \tag{1.19}$$

Using the relations $A(\alpha, \alpha) = 1$, $A_\beta(\alpha, \alpha) = 0$, we find that

$$E(\alpha, \alpha) = 0, \qquad E_\beta(\alpha, \alpha) = 0.$$

We represent $E(\alpha, \beta)$ by Taylor's formula, regarding $E(\alpha, \beta)$ as a function of β, and expanding $E(\alpha, \beta)$ about the point $\beta = \alpha$ with a remainder as a term of the second order. Thus

$$E(\alpha, \beta) = \tfrac{1}{2}(\alpha - \beta)^2 \int_0^1 E_{\beta\beta}[\alpha, \alpha + t(\beta - \alpha)]\, dt.$$

Upon setting

$$\tfrac{1}{2}(\alpha - \beta) \{\sin [\tfrac{1}{2}(\alpha - \beta)]\}^{-1} = M(\alpha, \beta) \qquad (\alpha \neq \beta)$$

with $M(\alpha, \alpha) = 1$, we see that $M(\alpha, \beta)$ is analytic on T, and that

$$S(g, \lambda) = \frac{1}{8\pi} \iint_T \left\{[g(\alpha) - g(\beta)]^2 M^2(\alpha, \beta) \int_0^1 E_{\beta\beta}[\alpha, \alpha + t(\beta - \alpha)]\, dt\right\} d\alpha\, d\beta. \tag{1.20}$$

It appears from (1.19) that $E_{\beta\beta}$ depends on λ and its first three derivatives. To two admissible functions λ, say λ and λ^*, we shall assign a distance $\lambda\lambda^*$ equal to the maximum value of

$$\begin{aligned}(1.21)\qquad &| \lambda(\theta) - \lambda^*(\theta) | + | \lambda'(\theta) - \lambda^{*\prime}(\theta) | + | \lambda''(\theta) - \lambda^{*\prime\prime}(\theta) | \\ &\qquad + | \lambda'''(\theta) - \lambda^{*\prime\prime\prime}(\theta) |.\end{aligned}$$

To two boundary vectors $g(\theta)$ and $g^*(\theta)$ we shall assign a distance gg^* equal to the maximum of the vector length

$$(1.22) \qquad | g(\theta) - g^*(\theta) |.$$

In the space of the sets (g, λ), distance will be defined by the sum

$$gg^* + \lambda\lambda^*.$$

A function of the sets (g, λ) will be regarded as continuous if continuous with respect to the metric of the sets (g, λ). Upon referring to (1.19), we see that $E_{\beta\beta}$ is continuous in (λ). Making use of (1.20), we obtain the following basic theorem.[14]

THEOREM 1.2. *The functional $S(g, \lambda)$ is continuous in its arguments, and when $A(g)$ is finite, equals the derivative $A_e(p)$ evaluated when $e = 0$.*

The functional $S(g, \lambda)$ is thus well defined and continuous even when $A(g)$ is infinite, and in many ways is more amenable to analysis than $A(g)$. The implications of this fact are very significant, and do not seem to have been sufficiently exploited.

2. **The fundamental theorem when** $n = 0$. We continue with the one-parameter family $H(u, v, e)$ of harmonic surfaces of §1, defined as in §1 by the (λ, e) "transform" $p(\theta, e)$ of $g(\theta)$. It will be convenient, however, to suppose that C is a circle with arbitrary radius a and center at the origin. Recall that $p(\theta, 0) \equiv g(\theta)$. We shall begin with the case where the components of $g(\theta)$ are analytic in θ. The vector $H(u, v, e)$ will be analytic on $\bar{B}$, and in particular will possess continuous partial derivatives on $\bar{B}$ of all orders. We have

$$A(p) = A(g, \lambda, e) = \frac{1}{2}\iint_B \left\{\left(\frac{\partial H}{\partial u}\right)^2 + \left(\frac{\partial H}{\partial v}\right)^2\right\} du\, dv.$$

If we differentiate under the integral sign with respect to e, integrate by parts in the classical way, and finally make use of the fact that H is harmonic, we find that

$$(2.1) \qquad A_e = \int_C \frac{\partial H}{\partial e}\left(\frac{\partial H}{\partial u}\, dv - \frac{\partial H}{\partial v}\, du\right).$$

[14] Cf. M. Morse and C. B. Tompkins, *The existence of minimal surfaces of general critical types*, Annals of Mathematics, vol. 40(1939), pp. 443–472.

In terms of the polar coördinates (r, θ), (2.1) yields the formula

$$A_e = \int_0^{2\pi} \frac{\partial H}{\partial e} \frac{\partial h}{\partial r} r \, d\theta, \qquad (r = a, e = 0), \tag{2.2}$$

where $h(u, v) = H(u, v, 0)$.

Recall that

$$p(\mu, e) \equiv g(\theta), \tag{2.3}$$

subject to the relation

$$\mu = \theta + e\lambda(\theta), \qquad | e | \leqq e(\lambda). \tag{2.4}$$

In (2.3) and (2.4) we regard e and μ as independent and θ as dependent, and differentiate with respect to e, holding μ fast. We then set $e = 0$. Thus

$$p_e(\mu, 0) = g'(\theta) \frac{\partial \theta}{\partial e},$$

$$0 = \frac{\partial \theta}{\partial e} + \lambda(\theta) \qquad (e = 0).$$

But

$$H(a \cos \mu, a \sin \mu, e) \equiv p(\mu, e),$$

so that

$$\left. \frac{\partial H}{\partial e} \right|^{e=0} = p_e(\mu, 0) = g'(\theta) \frac{\partial \theta}{\partial e} = -g'(\theta)\lambda(\theta) \qquad (r = a).$$

Finally

$$h(a \cos \theta, a \sin \theta) \equiv g(\theta),$$

so that when $r = a$

$$\frac{\partial h}{\partial \theta} = g'(\theta), \qquad \frac{\partial H}{\partial e} = -\frac{\partial h}{\partial \theta} \lambda(\theta) \qquad (e = 0).$$

Hence (2.2) takes the form

$$A_e = -\int_0^{2\pi} \lambda(\theta) \frac{\partial h}{\partial \theta} \frac{\partial h}{\partial r} r \, d\theta \qquad (r = a, e = 0). \tag{2.5}$$

In terms of the function $F(w)$ of the introduction, $h(u, v) = RF(w)$. Recall that

$$F'(w) = e^{-i\theta} \left(\frac{\partial h}{\partial r} - \frac{i}{r} \frac{\partial h}{\partial \theta} \right) \qquad (r \neq 0). \tag{2.6}$$

We introduce the "Weierstrass modul" $f(w) = F'(w) \cdot F'(w)$. Multiplying both sides of (2.6) by $w = re^{i\theta}$ and taking the scalar product of the respective members by themselves, we find that

$$w^2 f(w) = \left(r\frac{\partial h}{\partial r} - i\frac{\partial h}{\partial \theta}\right)^2. \tag{2.6$'$}$$

The imaginary parts of the respective members take the form

$$I[w^2 f(w)] = -2r\frac{\partial h}{\partial r}\frac{\partial h}{\partial \theta}.$$

Referring to (2.5), we have the following lemma.

LEMMA 2.1. *When the boundary vector $g(\theta)$ is analytic, the variational coefficient $A_e(g, \lambda, 0)$ takes the form*

$$A_e(g, \lambda, 0) = \frac{1}{2}\int_0^{2\pi} \lambda(\theta) I[w^2 f(w)]\, d\theta \qquad (r = a), \tag{2.7}$$

where $f(w)$ is the Weierstrass modul (0.8) belonging to the circular region bounded by C and to the boundary vector $g(\theta)$.

We turn to the case where $g(\theta)$ is merely continuous. Let $h^*(u, v)$ be the vector which is harmonic for $r < 1$, continuous for $r \leqq 1$, and such that

$$h^*(\cos\theta, \sin\theta) \equiv g(\theta).$$

Set

$$g_a(\theta) = h^*(a\cos\theta, a\sin\theta) \qquad (0 < a \leqq 1). \tag{2.8}$$

The vector $g_a(\theta)$ is analytic for $a < 1$. Let $z = \rho e^{\psi i}$ be an arbitrary complex point with $\rho < 1$. Hold z fast and set

$$\lambda_a(\theta) = \frac{1}{\pi}\frac{a^2 - \rho^2}{a^2 - 2a\rho\cos(\theta - \psi) + \rho^2} \qquad (\rho < a). \tag{2.9}$$

We have in $\lambda_a(\theta)$ the Poisson kernel for the circle $r = a$. In Lemma 2.1 take C as a circle $r = a < 1$. From (2.7) and the properties of the Poisson integral we find that

$$A_e(g_a, \lambda_a, 0) = I[z^2 f_a(z)],$$

where $f_a(z)$ is the Weierstrass modul belonging to the region $|z| < a$ and to the boundary vector $g_a(\theta)$. But $f_a(z) \equiv f_1(z)$, so that

$$A_e(g_a, \lambda_a, 0) = I[z^2 f_1(z)]. \tag{2.10}$$

For z fixed, the right member of (2.10) is independent of a, while the left member of (2.8) is continuous in its arguments by virtue of Theorem 1.2. Upon letting a tend to 1 as a limit, we obtain the following theorem.

THEOREM 2.1. *Corresponding to an arbitrary choice of the complex point* z *with* $|z| < 1$ *and choice of* λ *as the Poisson kernel* (2.9) *for the unit circle with pole* $\rho e^{\psi i} = z$,

$$A_e(g, \lambda, 0) = I[z^2 f(z)],$$

where $f(z)$ *is the Weierstrass modul belonging to the region* $|z| < 1$ *and to the boundary vector* $g(\theta)$.

The fundamental theorem of the introduction can now be established.

Suppose first that $A_e(g, \lambda, 0) = 0$ for every admissible λ. In particular λ can be chosen as in the preceding theorem. Hence $I[z^2 f(z)] = 0$ at each point $|z| < 1$. It follows that $z^2 f(z)$ is a constant. Upon setting $z = 0$, we see that this constant is null. Hence $f(z) \equiv 0$, and the harmonic surface $h^*(u, v)$ defined by g for $r \leqq 1$ is minimal.

Conversely, suppose that $h^*(u, v)$ is minimal. Then $f(z) \equiv 0$. We shall make use of (2.7), taking g as the function g_a of (2.8), $a < 1$, and recalling that $f_a(w) \equiv f(w)$. We infer that

$$A_e(g_a, \lambda, 0) = 0$$

for each admissible λ. Upon letting a tend to 1 and using Theorem 1.2, we infer that

$$A_e(g, \lambda, 0) = 0,$$

as required.

3. A series for $D(g)$. We return to the notation of the introduction. We suppose that C_0 is the unit circle and contains B in its interior. Let

$$(g_0, g_1, \cdots, g_n) = (g)$$

be the given boundary vectors, and $h(u, v)$ the harmonic surface thereby defined over B.

Let Γ_k be a circle on B concentric with C_k, with a radius ρ_k. Let R_k be the annular region bounded by C_k and Γ_k $(k = 0, \cdots, n)$. Set $w = u + iv$. The harmonic vector $h(u, v)$ is the real part of an analytic vector function $F(w)$. The function $F'(w)$ is single-valued and so admits a Laurent expansion on R_k in powers of $w - w_k$, where w_k is the center of C_k. From this expansion for $F'(w)$ we find that $F(w)$ has the form

$$F(w) = c_k \log (w - w_k) + \sum_m \gamma_k(m)(w - w_k)^m \qquad (m = 0, \pm 1, \pm 2, \cdots)$$

on R_k. Since $RF(w)$ is single-valued, we infer that c_k is real. We set[15]

$$\gamma_k(m) = A_k(m) - iB_k(m) \qquad (k = 0, \cdots, n),$$

[15] Cf. Douglas II, p. 338 for a similar procedure.

where $A_k(m)$ and $B_k(m)$ are real. Then on R_k,

$$(3.1)\quad \begin{aligned} h(u, v) = c_k \log r_k + A_k(0) &+ \sum_{m=1}^{\infty} r_k^m [A_k(m) \cos m\theta_k + B_k(m) \sin m\theta_k] \\ &+ \sum_{m=1}^{\infty} r_k^{-m}[A_k(-m) \cos m\theta_k - B_k(-m) \sin m\theta_k]. \end{aligned}$$

Let τ_k be taken between ρ_k and σ_k, the radius of C_k, and set

$$(3.2)\quad (\tau_0, \tau_1, \cdots, \tau_n) = (\tau).$$

Let $B(\tau)$ be a subregion of B in which for each k the boundary circle C_k of B is replaced by a concentric circle $C_k(\tau)$ of radius τ_k. Let

$$(3.3)\quad d(\tau) = \frac{1}{2} \iint_{B(\tau)} \left\{ \left(\frac{\partial h}{\partial u}\right)^2 + \left(\frac{\partial h}{\partial v}\right)^2 \right\} du\, dv,$$

so that

$$(3.4)\quad D(g) = \lim_{(\tau)=(\sigma)} d(\tau).$$

Upon integrating the right member of (3.3) by parts in the usual way and making use of the fact that $h(u, v)$ is harmonic, we find that

$$d(\tau) = \frac{1}{2} \sum_k \int_{C_k(\tau)} h \left(\frac{\partial h}{\partial u} dv - \frac{\partial h}{\partial v} du \right),$$

where the sense of integration on C_0 is counterclockwise, and along $C_1, \cdots, C_n$, clockwise. We introduce the multipliers

$$\epsilon_0 = 1, \qquad \epsilon_1 = \cdots = \epsilon_n = -1,$$

and write $d(\tau)$ in the form

$$d(\tau) = \frac{1}{2} \sum_k \int_0^{2\pi} \epsilon_k h \frac{\partial h}{\partial r_k} r_k \, d\theta_k \Big|^{r_k=\tau_k} \qquad (k = 0, \cdots, n),$$

$$(3.5)\quad d(\tau) = \frac{1}{4} \sum_k \epsilon_k \tau_k \frac{\partial}{\partial r_k} \int_0^{2\pi} h^2 \, d\theta_k \Big|^{r_k=\tau_k}.$$

To evaluate[15] the integrals in (3.5) we make use of Parseval's theorem, regarding (3.1) as a Fourier development with fixed r_k. For r_k between ρ_k and σ_k,

$$(3.6)\quad \begin{aligned} \int_0^{2\pi} h^2 \, d\theta_k = 2\pi [c_k \log r_k + A_k(0)]^2 \\ + \pi \sum_{m=1}^{\infty} \{ r_k^{2m} [A_k^2(m) + B_k^2(m)] + 2A_k(m) A_k(-m) \\ - 2B_k(m) B_k(-m) + r_k^{-2m} [A_k^2(-m) + B_k^2(-m)] \}. \end{aligned}$$

The right member of (3.6) will be shown to converge uniformly for r_k on any closed subinterval I_k of the values between ρ_k and σ_k exclusive. For the subseries of terms in (3.6) which involve *positive* powers of r_k converges for r_k between ρ_k and σ_k since one can apply Parseval's theorem to the sum of the corresponding terms in (3.1). Hence this subseries converges uniformly on I_k. Similarly, the subseries of (3.6) involving negative powers of r_k converges uniformly on I_k, and so (3.6) can be differentiated termwise with respect to r_k.

It follows from (3.5) and (3.6) that

$$\begin{aligned} d(\tau) &= \pi \sum_k \epsilon_k c_k [c_k \log \tau_k + A_k(0)] \\ &\quad + \pi \sum_m \tfrac{1}{2} m \sum_k \epsilon_k \{[A_k^2(m) + B_k^2(m)]\tau_k^{2m} - [A_k^2(-m) + B_k^2(-m)]\tau_k^{-2m}\} \end{aligned} \tag{3.7}$$

$$(k = 0, 1, \cdots, n;\ m = 1, 2, \cdots).$$

Consider the subseries $T(\tau)$ of (3.7) whose m-th term is

$$T(m, \tau) = \tfrac{1}{2}\pi m \left\{[A_0^2(m) + B_0^2(m)]\tau_0^{2m} + \sum_{k=1}^{n} [A_k^2(-m) + B_k^2(-m)]\tau_k^{-2m}\right\},$$

and the residual subseries,

$$\begin{aligned} S(\tau) &= \pi \sum_{k=0}^{n} \epsilon_k c_k [c_k \log \tau_k + A_k(0)] \\ &\quad - \pi \sum_m \tfrac{1}{2} m \left\{\sum_{k=1}^{n} [A_k^2(m) + B_k^2(m)]\tau_k^{2m} + [A_0^2(-m) + B_0^2(-m)]\tau_0^{-2m}\right\}. \end{aligned}$$

Let J_k be a closed interval of values of τ_k between σ_k and ρ_k which includes σ_k as an end point but which excludes ρ_k. We wish to show that $T(\tau)$ converges uniformly for τ_k on J_k. Now $S(\tau)$ converges uniformly on J_k, as one sees upon comparison with the convergent series Σ of corresponding terms in (3.1), or better, with the series obtained by applying Parseval's theorem to Σ. Hence

$$| S(\tau) | \leqq M \qquad (\tau_k \text{ on } J_k,\ k = 0, 1, \cdots, n), \tag{3.8}$$

where M is a constant independent of (τ). We note that $T(m, \tau) \geqq 0$, and conclude from (3.7) that

$$\sum_{m=1}^{\infty} T(m, \tau) \leqq d(\tau) + M \leqq D(g) + M. \tag{3.9}$$

From the continuity of $T(m, \tau)$ we infer that

$$\sum_{m=1}^{\infty} T(m, \sigma) \leqq D(g) + M. \tag{3.10}$$

But for τ_k on J_k, $T(m, \tau) \leqq T(m, \sigma)$. Hence the series (3.9) is termwise dominated by the convergent series (3.10) and so converges uniformly. The right

member of (3.7) is accordingly continuous at $(\tau) = (\sigma)$, and we conclude that

$$
\begin{aligned}
D(g) &= \pi \sum_k \epsilon_k c_k [c_k \log \sigma_k + A_k(0)] \\
&\quad + \pi \sum_m \tfrac{1}{2} m \sum_k \epsilon_k \{[A_k^2(m) + B_k^2(m)]\sigma_k^{2m} - [A_k^2(-m) + B_k^2(-m)]\sigma_k^{-2m}\} \qquad (3.11)
\end{aligned}
$$

$$(k = 0, 1, \cdots, n;\ m = 1, 2, \cdots;\ \epsilon_0 = 1,\ \epsilon_1 = \cdots = \epsilon_n = -1).$$

This formula is fundamental.

The formula when $n = 0$. In this case C_0 is the only circle. We take C_0 as the unit circle and let (r, θ) be polar coördinates with pole at the center of C_0. As is well known, the harmonic function $h(u, v)$ defined by the boundary vector $g(\theta)$ on C_0 has the form

$$h(u, v) = \tfrac{1}{2}a_0 + \sum_{m=1}^{\infty} r^m(a_m \cos m\theta + b_m \sin m\theta) \qquad (r < 1),$$

where (a_i, b_i) are Fourier vector coefficients of $g(\theta)$. Upon comparing this representation of $h(u, v)$ with the representation (3.1), we see that $c_0 = 0$, and that for positive values of m,

$$A_0(m) = a_m, \qquad B_0(m) = b_m, \qquad A_0(-m) = B_0(-m) = 0.$$

Hence (3.11) gives the formula

$$A(g) = \tfrac{1}{2}\pi \sum_m m(a_m^2 + b_m^2), \qquad (3.12)$$

as is well known.

4. **A second representation of** $D(g)$. We shall evaluate the coefficients in (3.1) in terms of the Fourier vector coefficients $a_k(m)$, $b_k(m)$ of $h(u, v)$ on C_k, and of the Fourier vector coefficients $\alpha_k(m)$, $\beta_k(m)$ of $h(u, v)$ on Γ_k. We multiply the respective members of (3.1) by $\cos m\theta_k$ and $\sin m\theta_k$, and integrate between 0 and 2π with r_k fixed between σ_k and ρ_k. By virtue of the continuity of the resulting integrals as functions of r_k, r_k can take on the values σ_k and ρ_k so that

$$(4.1) \qquad \begin{cases} 2[A_k(0) + c_k \log \sigma_k] = a_k(0), \\ 2[A_k(0) + c_k \log \rho_k] = \alpha_k(0), \end{cases}$$

$$(4.2) \qquad \begin{cases} A_k(m)\sigma_k^m + A_k(-m)\sigma_k^{-m} = a_k(m), \\ A_k(m)\rho_k^m + A_k(-m)\rho_k^{-m} = \alpha_k(m), \end{cases}$$

$$(4.3) \qquad \begin{cases} B_k(m)\sigma_k^m - B_k(-m)\sigma_k^{-m} = b_k(m), \\ B_k(m)\rho_k^m - B_k(-m)\rho_k^{-m} = \beta_k(m), \end{cases}$$

$$(k = 0, \cdots, n).$$

Set

$$t_0 = \frac{\rho_0}{\sigma_0}, \qquad t_h = \frac{\sigma_h}{\rho_h} \qquad (h = 1, \cdots, n).$$

From the preceding relations we see that

$$\begin{cases} 2c_h \log t_h = a_h(0) - \alpha_h(0), \\ 2c_0 \log t_0 = \alpha_0(0) - a_0(0). \end{cases} \qquad (h = 1, \cdots, n), \tag{4.4}$$

In (3.10) we set

$$\begin{aligned} A_h(-m)\sigma_h^{-m} &= a_h(m) - A_h(m)\sigma_h^m, \\ A_0(m)\sigma_0^m &= a_0(m) - A_0(-m)\sigma_0^{-m}, \\ B_h(-m)\sigma_h^{-m} &= -b_h(m) + B_h(m)\sigma_h^m, \\ B_0(m)\sigma_0^m &= b_0(m) + B_0(-m)\sigma_0^{-m}, \end{aligned} \qquad (h = 1, \cdots, n), \tag{4.5}$$

in accordance with (4.2) and (4.3). Making use of (4.1) and (4.4), we then reduce (3.11) to the form

$$\begin{aligned} D(g) = \frac{\pi}{4} \sum_k a_k(0) \frac{[\alpha_k(0) - a_k(0)]}{\log t_k} &+ \frac{\pi}{2} \sum_m m \sum_k [a_k^2(m) + b_k^2(m)] \\ &- \pi \sum_m m \sum_h \{a_h(m)A_h(m) + b_h(m)B_h(m)\}\sigma_h^m \\ &- \pi \sum_m m\{a_0(m)A_0(-m) - b_0(m)B_0(-m)\}\sigma_0^{-m} \\ &(h = 1, \cdots, n; k = 0, \cdots, n; m = 1, 2, \cdots). \end{aligned} \tag{4.6}$$

From (4.2) and (4.3) we see that, for $h = 1, \cdots, n$,

$$\begin{cases} A_h(m)\sigma_h^m = t_h^m \dfrac{[\alpha_h(m) - t_h^m a_h(m)]}{1 - t_h^{2m}}, \\ A_0(-m)\sigma_0^{-m} = t_0^m \dfrac{[\alpha_0(m) - t_0^m a_0(m)]}{1 - t_0^{2m}}, \end{cases} \tag{4.7}$$

$$\begin{cases} B_h(m)\sigma_h^m = t_h^m \dfrac{[\beta_h(m) - t_h^m b_h(m)]}{1 - t_h^{2m}}, \\ -B_0(-m)\sigma_0^{-m} = t_0^m \dfrac{[\beta_0(m) - t_0^m b_0(m)]}{1 - t_0^{2m}}. \end{cases} \tag{4.8}$$

Upon using (3.12), (4.6), (4.7), and (4.8), we find that

$$D(g) = A(g_0) + \cdots + A(g_n) + R(g), \tag{4.9}$$

where

$$R(g) = T(g) + R_0(g) + \cdots + R_n(g), \tag{4.10$'$}$$

$$T(g) = \frac{\pi}{4} \sum_k a_k(0) \frac{[\alpha_k(0) - a_k(0)]}{\log t_k} \qquad (k = 0, \cdots, n), \tag{4.10$''$}$$

$$(4.10)''' \quad R_k(g) = \pi \sum_m \left[mt_k^m a_k(m) \frac{[t_k^m a_k(m) - \alpha_k(m)]}{1 - t_k^{2m}} + mt_k^m b_k(m) \frac{[t_k^m b_k(m) - \beta_k(m)]}{1 - t_k^{2m}} \right]$$

$$(k = 0, \cdots, n; m = 1, 2, \cdots).$$

Let M and μ be constants such that $M > 1$, $\mu < 1$, and suppose that

$$(4.11) \qquad | g_k | < \tfrac{1}{2}M, \qquad t_k < \mu < 1 \qquad (k = 0, 1, \cdots, n).$$

Then

$$(4.12) \qquad 1 - t_k^{2m} > 1 - \mu,$$

and the Fourier coefficients $a_k(m)$, $\alpha_k(m)$, etc. have magnitudes at most M. Moreover,

$$(4.13) \qquad | t_k^m a_k(m) - \alpha_k(m) | < 2M, \qquad | t_k^m b_k(m) - \beta_k(m) | < 2M,$$

so that the m-th term in R_k is in magnitude at most

$$(4.14) \qquad \frac{4\pi M^2 m\mu^m}{1 - \mu} \qquad (\mu < 1).$$

The series of terms (4.14) is convergent so that the series R_k converges absolutely. This fact justifies the regrouping of the terms of (3.11) in the form (4.9).

We shall examine the dependence of R_k on the boundary vectors (g) and the circles (C). The distance $d(gg^*)$ between two admissible sets (g) and (g^*), both given in terms of the angles (θ), will be defined as the sum

$$g_0 g_0^* + g_1 g_1^* + \cdots + g_n g_n^*,$$

where $g_k g_k^*$ has been defined in (1.22). The circles (C) depend on the circle parameters (η) of the introduction. The distance $d(\eta\eta^*)$ between two admissible sets of circle parameters (η) and (η^*) will be taken as the Euclidean distance between (η) and (η^*), regarded as points in the space of their coördinates. Finally the distance between two admissible sets (g, η) and (g^*, η^*) will be taken as the sum

$$d(gg^*) + d(\eta\eta^*).$$

The sets (g, η) may then be regarded as points on an abstract metric space. The concept of continuity of a function $\varphi(g, \eta)$ will be referred to the general notion of continuity of a function on an abstract metric space.

The coefficients $a_k(m)$, $b_k(m)$ are continuous in (g), but independent of (η). The coefficients $\alpha_k(m)$ and $\beta_k(m)$ are continuous in (g, η) at least as long as the circles (Γ) remain fixed and interior to B. For under such circumstances $h(u, v)$, evaluated on (Γ), depends continuously on (g, η). If (4.11) holds for a given set (g, η), it will hold for a sufficiently small neighborhood Σ of (g, η) so

that the m-th term in R_k will still be dominated on Σ by the term (4.14). Thus R_k will converge uniformly on Σ.

Hence $R_k(g)$ varies continuously with (g, η), (g, η) remaining admissible. The same is clearly true of $T(g)$, and hence of $R(g)$.

Under circumstances to be described in the next section, the vectors (g), the circles (C), and the ratios t_k will vary with a parameter e in such a way that

$$\left|\frac{\partial a_k(m)}{\partial e}\right| < mN, \quad \left|\frac{\partial b_k(m)}{\partial e}\right| < mN, \quad \left|\frac{\partial \alpha_k(m)}{\partial e}\right| < mN, \quad \left|\frac{\partial \beta_k(m)}{\partial e}\right| < mN, \tag{4.15}$$

$$\left|\frac{\partial t_k}{\partial e}\right| < N \qquad (m = 1, 2, \cdots; k = 0, 1, \cdots, n), \tag{4.16}$$

where N is a constant > 1, independent of e, k, and m. When (4.11), (4.15), and (4.16) hold, the e-derivative of the m-th term of R_k is in magnitude at most

$$32\pi N M^2 \frac{m^2 \mu^{m-1}}{(1 - \mu)^2},$$

as one sees upon first verifying that

$$\left|\frac{\partial}{\partial e}\left(\frac{t_k^m}{1 - t_k^{2m}}\right)\right| < \frac{2mN\mu^{m-1}}{(1 - \mu)^2}.$$

Hence the series of e-derivatives of the terms of R_k converges uniformly, and we have the following fundamental lemma.

LEMMA 4.1. *When (4.11), (4.15), and (4.16) hold, the remainder R in (4.9) has an e-derivative which is continuous in e and in any other arguments upon which the Fourier coefficients $a_k(m)$, $\alpha_k(m)$, etc., and the ratios t_k depend continuously with their e-derivatives.*[16]

The remainder $R(g)$ is continuous in (g, η). If we define $D(g)$ as in (3.4) and use (4.9) to represent the approximation $d(\tau)$, we arrive at the following theorem.

THEOREM 4.1. *A necessary and sufficient condition that $D(g)$ be finite is that $A(g_0), \cdots, A(g_n)$ be finite.*

The case of two contours, $n = 1$. Formulas (4.10) can be considerably simplified in case $n = 1$. In fact (4.10) remains valid if Γ_0 and Γ_1 are replaced by C_0 and C_1 respectively. This amounts to letting ρ_1 tend to σ_0 and ρ_0 tend to σ_1 as a limit. The coefficients $\alpha_1(m)$ and $\alpha_0(m)$ will then tend to $a_0(m)$ and $a_1(m)$ respectively. Similarly, $\beta_1(m)$ and $\beta_0(m)$ tend to $b_0(m)$ and $b_1(m)$ respectively. The ratios t_0 and t_1 will have a common limit t given by σ_1/σ_0. The series $R_k(g)$ given by (4.10)''' is dominated termwise at all stages of this limiting process

[16] The parameter e is not to be confused with the set of constants (e) appearing as arguments of the function $D(g, \eta, \lambda, e)$. However, we shall have occasion to set $e = e_k$.

by a series of constant terms of the form (4.14). We are accordingly justified in setting

$$t_0 = t_1 = t,$$

$$\alpha_1(m) = a_0(m), \qquad \alpha_0(m) = a_1(m),$$

$$\beta_1(m) = b_0(m), \qquad \beta_0(m) = b_1(m)$$

in (4.10). We thereby find (cf. Douglas II, loc. cit., p. 339) that

$$D(g) = A(g_0) + A(g_1) + T(g) + R^*(g), \tag{4.17}$$

where

$$T(g) = -\frac{\pi}{4}\frac{[a_1(0) - a_0(0)]^2}{\log t},$$

$$R^*(g) = \pi \sum_{m=1}^{\infty} mt^{2m} \frac{[a_0^2(m) + b_0^2(m) + a_1^2(m) + b_1^2(m)]}{1 - t^{2m}}$$

$$- 2\pi \sum_{m=1}^{\infty} mt^m \frac{[a_0(m)a_1(m) + b_0(m)b_1(m)]}{1 - t^{2m}}.$$

5. **The first variation of D.** Recall that $\lambda_k(\theta_k)$ is an analytic function of θ_k with a period 2π. The set $(\lambda_0, \cdots, \lambda_n)$ has been denoted by (λ). The distance between two sets (λ) and (λ^*) will be taken as the sum

$$d(\lambda_0\lambda_0^*) + \cdots + d(\lambda_n\lambda_n^*), \tag{5.1}$$

where $d(\lambda_k\lambda_k^*)$ is defined in (1.21). The distance between sets (g, η, λ, e) and sets $(g^*, \eta^*, \lambda^*, e^*)$ will be taken as

$$d(gg^*) + d(\eta\eta^*) + d(\lambda\lambda^*) + d(ee^*), \tag{5.2}$$

where $d(ee^*) = |e - e^*|$. The concept of continuity of functions of sets (g, η, λ, e) will then be referred to the general concept of continuity of a point on an abstract metric space. Continuity of functions of sets (g, η, λ) or (g, η, e) will be similarly defined.

We shall make a study of the function $D(g, \eta, \lambda, e)$ of (0.6) and of its derivatives

$$D_{e_k}(g, \eta, \lambda, e), \qquad D_{\eta_i}(g, \eta, \lambda, e),$$

showing that these derivatives are continuous in their arguments at admissible sets (g, η, λ, e) at which D is finite.

We shall begin with a study of the Fourier coefficients $a_k(m)$, $b_k(m)$ of the (λ_k, e_k) transform $p(\theta_k, e_k)$ of g_k. These coefficients will be functions of the sets (g, λ, e) as well as of k and m, but will be independent of (η). In particular for $m > 0$,

$$a_k(m) = \frac{1}{\pi}\int_0^{2\pi} p_k(\mu, e_k) \cos m\mu \, d\mu. \tag{5.3}$$

Upon setting $\mu = \theta + e_k\lambda_k(\theta)$ subject to the condition (0.3) on e_k, we see that

$$(5.4) \qquad a_k(m) = \frac{1}{\pi}\int_0^{2\pi} g_k(\theta)\cos[m\theta + e_k m\lambda_k(\theta)][1 + e_k\lambda_k'(\theta)]\,d\theta.$$

We see from (5.4) that $a_k(m)$ and its e_j-derivatives are continuous in (g, λ, e). A similar conclusion holds for $b_k(m)$.

We turn to the Fourier coefficients $\alpha_k(m)$ and $\beta_k(m)$. Let $G(x, y, u, v, \eta)$ be the Green's function with pole at the point (u, v), set up for the region B^η bounded by the respective circles C_k^η with circle parameters (η). The point (x, y) shall vary near (C^η) with $(x, y) \neq (u, v)$. With respect to the variables (x, y), $\partial G/\partial N$ shall represent the inner normal derivative of G, evaluated near (C^η). The partial derivatives

$$(5.5) \qquad \frac{\partial^2 G}{\partial x\,\partial N}, \qquad \frac{\partial^2 G}{\partial y\,\partial N}, \qquad \frac{\partial^2 G}{\partial \eta_i\,\partial N} \qquad (i = 1, \cdots, 3n)$$

exist and are continuous in their arguments for (x, y) near (C^η), and (u, v) interior to B^η, $(x, y) \neq (u, v)$. This can be established by an elementary analysis of the properties of G. The vector $H(u, v, e)$ is given by the formula[17]

$$(5.6) \quad H(u, v, e) = \frac{1}{2\pi}\sum_k \int_{C_k^\eta} p_k(\mu_k, e_k)\frac{\partial}{\partial N}G(x, y, u, v, \eta)\rho_k\,d\mu_k \quad (k = 0, 1, \cdots, n),$$

where on C_k^η

$$(5.7) \qquad x = u_k + \sigma_k\cos\mu_k, \qquad y = v_k + \sigma_k\sin\mu_k.$$

We set

$$(5.8) \qquad \mu_k = \theta_k + e_k\lambda_k(\theta_k)$$

in (5.6), subject to (0.3), so that $H(u, v, e)$ takes the form

$$(5.9) \qquad H(u, v, e) = \frac{1}{2\pi}\sum_k \int_{C_k^\eta} g_k(\theta)[1 + e_k\lambda_k'(\theta)]L_k\,d\theta_k,$$

where, subject to (5.7) and to (5.8),

$$\rho_k\frac{\partial G}{\partial N} = L_k(e_k, \theta_k, u, v, \eta) \qquad [(x, y) \text{ on } C_k^\eta;\ (u, v) \text{ on } B^\eta].$$

By virtue of the existence and continuity of the partial derivatives (5.5), L_k is continuously differentiable in the arguments e_k and η_i. We can accordingly obtain the e_k-derivative or η_i-derivative of $H(u, v, e)$ by differentiating under the integral sign in (5.9). Recalling that $\alpha_k(m)$ and $\beta_k(m)$ are Fourier coefficients of H taken on the respective circles Γ_k, we conclude that the functions $\alpha_k(m)$ and $\beta_k(m)$ as well as their e_j- and η_i-derivatives are continuous in (g, η, λ, e).

We shall use Lemma 4.1 to prove a lemma concerning $R_k(p)$, where (p) is the (λ, e) transform of (g). We write $R_k(p)$ as a function $R_k(g, \eta, \lambda, e)$.

[17] The sense of integration is such that μ_0 varies from 0 to 2π, and $\mu_1, \cdots, \mu_n$ vary from 2π to 0.

LEMMA 5.1. *The e_j-derivative of $R_k(g, \eta, \lambda, e)$ exists and is continuous at each admissible set $(g^0, \eta^0, \lambda^0, e^0)$.*

Corresponding to the given set $(g^0, \eta^0, \lambda^0, e^0)$, there exists a neighborhood Σ of $(g^0, \eta^0, \lambda^0, e^0)$ so small that on Σ, for each k,

$$(5.10) \qquad | g_k(\theta) | < \tfrac{1}{2}M, \qquad | \lambda_k(\theta) | < \Lambda, \qquad | \lambda_k'(\theta) | < \Lambda,$$

where M and Λ are constants independent of k and θ. It will be convenient to suppose M and Λ at least 1. We see that $| a_k(m) |$ and $| b_k(m) |$ are less than M. From (5.4) and (5.10) we infer that

$$(5.11) \qquad \left| \frac{\partial a_k(m)}{\partial e_k} \right| < 4mM\Lambda^2, \qquad \left| \frac{\partial b_k(m)}{\partial e_k} \right| < 4mM\Lambda^2,$$

making use of the fact that $| e_k | < 1$.

We seek limitations similar to (5.11) on the e_j-derivatives of $\alpha_k(m)$ and $\beta_k(m)$. To that end we shall differentiate under the integral sign in (5.9). If the circles (Γ) are fixed and the neighborhood Σ of $(g^0, \eta^0, \lambda^0, e^0)$ is sufficiently small, the circles (C^η) will be bounded from the circles (Γ), so that the point (u, v) on (Γ) will be bounded from the point (x, y) on (C^η). Hence the e_j-derivative of L_k will have an absolute value which is bounded on Σ. Hence $| H_{e_j} |$ will admit a bound A for (u, v) on (Γ) independent of (g, η, λ, e) on Σ. Upon differentiating H under the integral sign in the Fourier integrals giving $\alpha_k(m)$ and $\beta_k(m)$, we infer that

$$(5.12) \qquad \left| \frac{\partial \alpha_k(m)}{\partial e_j} \right| \leqq 2A, \qquad \left| \frac{\partial \beta_k(m)}{\partial e_j} \right| \leqq 2A \qquad (k, j = 0, 1, \cdots, n).$$

We can apply Lemma 4.1. For we have seen that the coefficients

$$a_k(m), \qquad b_k(m), \qquad \alpha_k(m), \qquad \beta_k(m)$$

and their e_j-derivatives are continuous in (g, η, λ, e), the circles (Γ) remaining fixed, and (g, η, λ, e) admissible. In Lemma 4.1 set $e = e_j$. Conditions (4.11) and (4.15) of Lemma 4.1 are satisfied for (g, η, λ, e) on Σ, by virtue of (5.10), (5.11) and (5.12). Condition (4.16) of Lemma 4.1 holds since t_k is independent of e_j. Hence $R_k(p)$ has an e_j-derivative which is continuous on Σ, and the proof of the lemma is complete.

We continue with a proof of Lemma 5.2 again using Lemma 4.1.

LEMMA 5.2. *The η_i-derivative of $R_k(g, \eta, \lambda, e)$ exists and is continuous at each admissible set $(g^0, \eta^0, \lambda^0, e^0)$.*

We choose a neighborhood Σ of $(g^0, \eta^0, \lambda^0, e^0)$ as in the preceding lemma. Relations (5.10) then hold. The coefficients $a_k(m)$, $b_k(m)$ are independent of (η). The coefficients $\alpha_k(m)$, $\beta_k(m)$ and their η_i-derivatives are the Fourier coefficients respectively of $H(u, v, e)$ and $H_{\eta_i}(u, v, e)$ evaluated on Γ_k. The function $H(u, v, e)$ is given by the integral (5.9), and $H_{\eta_i}(u, v, e)$ can be obtained by differentiating (5.9) under the integral sign. For sets (g, η, λ, e) on Σ the

circles (C^η) are bounded from the circles (Γ), so that the point (u, v) on (Γ) will be bounded from the point (x, y) on (C^η). With the variables so restricted, the η_i-derivative of L_k in (5.9) will be bounded in absolute value so that $| H_{\eta_i}(u, v, e) |$ will be bounded by a constant B. The derivatives

$$\left| \frac{\partial \alpha_k(m)}{\partial \eta_i} \right|, \qquad \left| \frac{\partial \beta_k(m)}{\partial \eta_i} \right| \qquad [(g, \eta, \lambda, e) \text{ on } \Sigma]$$

will be bounded by $2B$ independently of k and i. Finally if Σ is sufficiently small, the ratios

$$t_0 = \frac{\rho_0}{\sigma_0}, \qquad t_h = \frac{\sigma_h}{\rho_h} \qquad (h = 1, \cdots, n)$$

will be bounded and will have bounded η_i-derivatives, since the denominators σ_0 and ρ_h will be bounded from zero.

We can thus apply Lemma 4.1, setting e in Lemma 4.1 equal to η_i, and infer that the η_i-derivative of $R_k(g, \eta, \lambda, e)$ exists and is continuous at each admissible set $(g^0, \eta^0, \lambda^0, e^0)$.

We complete these results in the fundamental theorem.

THEOREM 5.1. *The variational coefficients D_{e_k} and D_{η_i} exist and are continuous in the arguments (g, η, λ, e) at each admissible set of these arguments for which D is finite.*

Let (p) be the (λ, e) transform of (g), defined on the circles (C^η). It follows from (4.9) and (4.10) that

$$D(p) = A(p_0) + \cdots + A(p_n) + T(p) + R_0(p) + \cdots + R_n(p). \tag{5.13}$$

We have seen in Theorem 1.2 that $A_{e_k}(p_k)$ is continuous in its arguments (g_k, λ_k, e_k), and that $A(p_k)$ does not depend on (η). According to Lemmas 5.1 and 5.2, $R_k(p)$ has continuous e_k- and η_i-derivatives. The same is true of $T(p)$, as follows at once from the continuity and differentiability of the coefficients a_k, α_k, etc., and the ratios t_k.

The theorem follows from (5.13).

6. **The fundamental theorem.** We begin this section with a lemma.

Let $\psi(r, \theta)$ be a scalar function which is harmonic on an annular region of the (u, v)-plane of the form $r_0 \leqq r < 1$. Let (ρ, φ) be polar coördinates of a point for which $r_0 \leqq \rho < 1$, and suppose that $\rho < r$. Let $P(r, \theta, \rho, \varphi)$ be the Poisson kernel for the disc $u^2 + v^2 < r^2$ with pole at the point (ρ, φ). Set

$$\int_0^{2\pi} P(r, \theta, \rho, \varphi)\psi(r, \theta)\, d\theta = \zeta^{(r)}(\rho, \varphi). \tag{6.1}$$

Regarded as a function of (ρ, φ), $\zeta^{(r)}(\rho, \varphi)$ is harmonic for $\rho < r$ and, if extended by the definition

$$\zeta^{(r)}(r, \varphi) = \psi(r, \varphi),$$

is continuous for $\rho \leqq r$.

LEMMA 6.1.[18] *If for each fixed* $\rho < 1$

$$\lim_{r=1} \zeta^{(r)}(\rho, \varphi) = 0 \tag{6.2}$$

uniformly in φ, $\psi(r, \theta)$ *tends to* 0 *uniformly with respect to* θ *as* r *tends to* 1.

The difference

$$\psi(\rho, \varphi) - \zeta^{(r)}(\rho, \varphi) = M^{(r)}(\rho, \varphi) \qquad (r_0 < r < 1) \tag{6.3}$$

is harmonic in the variables (ρ, φ) for $r_0 \leqq \rho < r$ and vanishes for $\rho = r$. Hence $M^{(r)}(\rho, \varphi)$ can be continued by reflection so as to be harmonic over the ring $\Sigma(r)$,

$$r_0 \leqq \rho \leqq \frac{r^2}{r_0}.$$

The maximum m_r of $| M^{(r)}(\rho, \varphi) |$ on $\Sigma(r)$ will be assumed when $r = r_0$. It is the maximum of

$$| \psi(r_0, \varphi) - \zeta^{(r)}(r_0, \varphi) |$$

with respect to φ and tends to the maximum of $\psi(r_0, \varphi)$ as r tends to 1, by virtue of (6.2).

Let r_1 be a constant such that $r_0 < r_1 < 1$. For $r_1 \leqq r < 1$, m_r will admit a bound N independent of r. Suppose r_1 so near 1 that $r_1^2 > r_0$. Then $\Sigma(r_1)$ will contain the unit circle in its interior. Let S be a ring interior to $\Sigma(r_1)$, concentric with $\Sigma(r_1)$, containing the unit circle in its interior. Let d be the distance of the boundary of S from that of $\Sigma(r_1)$. Then on S

$$| M_\rho^{(r)}(\rho, \varphi) | < \frac{2N}{d}, \tag{6.4}$$

in accordance with the theory of harmonic functions.

It follows from (6.4) and from the fact that $M^{(r)}(r, \varphi) = 0$ that on S

$$| \psi(\rho, \varphi) - \zeta^{(r)}(\rho, \varphi) | \leqq \frac{2N}{d} (r - \rho).$$

We hold (ρ, φ) fast and let r tend to 1. By virtue of (6.2), $\zeta^{(r)}(\rho, \varphi)$ tends to 0 so that

$$| \psi(\rho, \varphi) | \leqq \frac{2N}{d} (1 - \rho).$$

The proof of the lemma is complete.

The lemma admits various extensions. It is clear that the unit circle can be replaced by an arbitrary circle C, and that the annular region adjoining C may be either interior or exterior to C. Further, under the hypothesis of the lemma, $\psi(\rho, \varphi)$ can be harmonically continued over C so as to be zero on C.

[18] Cf. Courant, loc. cit., p. 712.

We return to a study of the Dirichlet sum $D(g, \eta, \lambda, e)$ of (0.6). Corresponding to the relation (2.5) we here have the relation

$$D_{e_k}(g, \eta, \lambda, 0) \equiv -\int_{C_k^\eta} \lambda_k(\theta_k) \frac{\partial h}{\partial \theta_k} \frac{\partial h}{\partial r_k} \sigma_k \, d\theta_k .$$

When the boundary vectors (g) are analytic, $D_{e_k}(g, \eta, \lambda, 0)$ takes the form

$$D_{e_k}(g, \eta, \lambda, 0) = \frac{1}{2} \int_{C_k^\eta} \lambda_k(\theta_k) I[(w - w_k)^2 f(w)] \, d\theta_k , \tag{6.5}$$

where C_0^η is to be taken counterclockwise, and $C_1^\eta, \cdots, C_n^\eta$ clockwise, essentially as in Lemma 2.1.

We can now prove the following theorem.

THEOREM 6.1. *Corresponding to the boundary vectors (g) and circle parameters (η), let $f(w)$ be the Weierstrass modul of the harmonic function $h(u, v)$ determined by (g). A necessary and sufficient condition that $D_{e_k}(g, \eta, \lambda, 0) = 0$ for every admissible set (λ) and for the given set (g, η) is that the function*

$$I[(w - w_k)^2 f(w)] \qquad (k = 0, 1, \cdots, n) \tag{6.6}$$

admit a harmonic continuation over C_k^η, and in particular be null on C_k^η.

Let $[C(\tau)]$ be a set of circles interior to B^η of radii $(\tau_0, \cdots, \tau_n) = (\tau)$, concentric with the respective circles (C^η), and so near (C^η) that they bound a connected region $B(\tau)$. Let (g^τ) be the boundary vectors defined by $h(u, v)$ on the respective circles $[C(\tau)]$ in terms of the respective angles θ_k.

Let (η^τ) be the circle parameters determined by the circles $[C(\tau)]$. Referring to (6.1), let

$$\lambda_k^\tau(\theta_k) = P(\tau_k, \theta_k, \rho_k, \varphi_k) \qquad (\rho_k < \tau_k \leqq \sigma_k) \tag{6.7}$$

holding (ρ_k, φ_k) fast. For $\tau_k < \sigma_k$, the vectors (g^τ) are analytic in the parameters (θ) so that (6.5) is applicable. Hence for $w - w_k = \tau_k e^{\theta_k i}$

$$D_{e_k}(g^\tau, \eta^\tau, \lambda^\tau, 0) = \frac{1}{2} \int_{C_k(\tau)} \lambda_k^\tau(\theta_k) I[(w - w_k)^2 f(w)] \, d\theta_k \qquad (\tau_k < \sigma_k). \tag{6.8}$$

Set $(\lambda^\sigma) = (\lambda)$. If (τ) tends to (σ), $(g^\tau, \eta^\tau, \lambda^\tau)$ tends to (g, η, λ). By virtue of Theorem 5.1 the left member of (6.8) then tends to $D_{e_k}(g, \eta, \lambda, 0)$. Moreover, this approach is uniform with respect to the parameter φ_k of (6.7). For if ρ_k is fixed, if (τ) ranges on the closure of a sufficiently small neighborhood of (σ) and φ_k is arbitrary, the sets $(g^\tau, \eta^\tau, \lambda^\tau)$ form a compact ensemble on which the left member of (6.8) is continuous.

Under the hypothesis that $D_{e_k}(g, \eta, \lambda, 0) = 0$, we conclude that, for each k,

$$\lim_{(\tau)=(\sigma)} \int_{C_k(\tau)} \lambda_k^\tau(\theta_k) I[(w - w_k)^2 f(w)] \, d\theta_k = 0$$

uniformly with respect to the parameter φ_k of (6.7). It follows from Lemma 6.1 and its extensions that the function (6.6), harmonically continued, is null on C_k^η as stated, and we conclude that the condition of the theorem is necessary.

To prove the condition of the theorem sufficient we start with relation (6.8), assuming that the function (6.6) on $C_k(\tau)$ tends to zero uniformly as (τ) tends to (σ). By virtue of the continuity of the left member of (6.8), we can conclude that $D_{e_k}(g, \eta, \lambda, 0) = 0$, and the proof of the theorem is complete.

Formulas[19] *for* $D_{\eta_i}(g, \eta, \lambda, 0)$. We shall hold (g) fast and begin by varying the radii τ_k of the circles C_k, denoting these circles as previously by $C_k(\tau)$ and holding the centers $w_k = u_k + iv_k$ fast. The initial value of (τ) is denoted by (σ). As previously let (r_k, θ_k) be polar coördinates with pole at the point w_k. Let[20] $M(r_k, \theta_k, \tau)$ represent the harmonic surface determined by (g) on the region $B(\tau)$ bounded by the circles $[C(\tau)]$. Let $d(g, \tau)$ be the Dirichlet sum over $B(\tau)$ determined by $M(r_k, \theta_k, \tau)$. We shall begin by supposing the boundary vectors (g) analytic, and shall obtain a formula for $d_{\tau_k}(g, \sigma)$.

The integral $d(g, \tau)$ depends on τ_k both in its limits and in its integrand, so that d_{τ_k} consists of two principal terms. Indicating evaluation when $(\tau) = (\sigma)$ by a superscript 0, we have

$$(6.9)\quad d^0_{\tau_k} = -\frac{1}{2}\int_{C_k(\sigma)}\left[M^{02}_{r_k} + \frac{1}{\sigma_k^2}M^{02}_{\theta_k}\right]ds - \int_{C_k(\sigma)} M^0_{r_k}M^0_{\tau_k}\,ds \quad (k = 1, 2, \cdots, n).$$

Of these integrals the first arises from the fact that the limits depend on τ_k, while the second arises upon differentiating $d(g, \tau)$ under the integral sign and integrating by parts in the usual way.

Upon differentiating the relation

$$M(\tau_k, \theta_k, \tau) \equiv g_k(\theta_k)$$

with respect to τ_k, we find that

$$M_{r_k} + M_{\tau_k} = 0 \qquad (r_k = \tau_k),$$

so that (6.9) takes the form

$$d^0_{\tau_k} = \frac{1}{2}\int_{C_k(\sigma)}\left[M^{02}_{r_k} - \frac{M^{02}_{\theta_k}}{\sigma_k^2}\right]ds.$$

The relation (2.6)′ is here replaced by the relation

$$(w - w_k)^2 f(w) = (r_k M^0_{r_k} - iM^0_{\theta_k})^2,$$

from which it follows that

$$(6.10)\qquad d^0_{\tau_k} = \frac{1}{2}\int_{C_k(\sigma)} R[(w - w_k)^2 f(w)]\,\frac{ds}{\sigma_k^2}.$$

[19] Cf. Courant, loc. cit., p. 714.

[20] $M(\tau_k, \theta_k, \tau)$ is a vector in the space $(x^1, \cdots, x^m)$. This function depends upon k, but we have omitted a subscript k for the sake of simplicity of notation.

We shall show that (6.10) holds even when the boundary vectors are merely continuous, whenever the Weierstrass modul $f(w)$ is analytically continuable over the circles $C_k(\sigma)$. To establish this we use the boundary vectors (g^τ) and circles $[C(\tau)]$ of the proof of Theorem 6.1. For $(\tau) \neq (\sigma)$, the vectors (g^τ) are analytic in the angles (θ) and (6.10) yields the relation

$$d_{\tau_k}(g^\tau, \tau) = \frac{1}{2}\int_{C_k(\tau)} R[(w - w_k)^2 f(w)] \frac{ds}{\tau_k^2} \qquad (k = 1, \cdots, n). \tag{6.11}$$

When $f(w)$ is analytically continuable over $C_k(\sigma)$, the right member of (6.11) is continuous in (τ), even when $(\tau) = (\sigma)$. The left member of (6.11) is likewise continuous in (τ) by virtue of Theorem 6.1. It follows that (6.11) holds when $(\tau) = (\sigma)$, that is, (6.10) holds when the vectors (g) are merely continuous in (θ).

We shall now vary the coördinates (u, v) of the centers of the circles C_k^η $(k = 1, \cdots, n)$, denoting the variable coördinates by (a_k, b_k) and the initial coördinates by (u_k, v_k) as previously. We hold the radii of the circles C_k^η fast as well as the boundary vectors (g). Let $N(u, v, a, b)$ be the harmonic vector determined by (g) on the region $B(a, b)$ bounded by the circles with centers (a_k, b_k). Let $\delta(g, a, b)$ be the Dirichlet sum over $B(a, b)$ determined by $N(u, v, a, b)$. Evaluation when $(a_k, b_k) = (u_k, v_k)$ $(k = 1, \cdots, n)$ will be indicated by adding the superscript 0. In the case where the vectors $g_k(\theta_k)$ are analytic we find that

$$\delta^0_{a_k} = -\int_{C_k^\eta} N^0_{a_k} \frac{\partial N^0}{\partial r_k} ds + \frac{1}{2}\int_{C_k^\eta} [N_u^{0^2} + N_v^{0^2}]\, dv \qquad (k = 1, 2, \cdots, n), \tag{6.12}$$

the first integral arising from differentiating $\delta(g, a, b)$ under the integral sign and integrating by parts, and the second integral arising from the way in which a_k enters into the limits. The sense of integration on C_k^η is clockwise.

From the identity

$$N(a_k + \sigma_k \cos\theta_k,\, b_k + \sigma_k \sin\theta_k,\, a, b) = g_k(\theta_k),$$

we infer, by differentiating with respect to a_k, that on C_k^η,

$$N^0_u + N^0_{a_k} = 0,$$

so that (dropping η from C_k^η for simplicity of notation)

$$\delta^0_{a_k} = \int_{C_k} N^0_u \frac{\partial N^0}{\partial r_k} ds - \frac{1}{2}\int_{C_k} [N_u^{0^2} + N_v^{0^2}] \cos\theta_k\, ds. \tag{6.13}$$

Upon recalling that

$$\frac{\partial N}{\partial r_k} = N_u \cos\theta_k + N_v \sin\theta_k,$$

we reduce (6.13) to the form

$$\delta^0_{a_k} = \int_{C_k} N^0_u N^0_v \sin\theta_k\, ds + \frac{1}{2}\int_{C_k} [N^{0^2}_u - N^{0^2}_v] \cos\theta_k\, ds,$$

$$\delta^0_{a_k} = \frac{1}{2\sigma_k}\int_{C_k} R[(w - w_k)f(w)]\, ds \qquad (k = 1, 2, \cdots, n). \tag{6.14}$$

Similarly one shows that when the vectors (g) are analytic in the variables (θ),

$$\delta^0_{b_k} = \frac{1}{2\sigma_k}\int_{C_k} I[(w_k - w)f(w)]\, ds \qquad (k = 1, \cdots, n). \tag{6.15}$$

When the Weierstrass modul $f(w)$ is analytically continuable over the circles (C^η), it follows, as in the proof of (6.10), that (6.14) and (6.15) hold even if (g) is merely continuous in the variables (θ). We thus have the following lemma.

LEMMA 6.2. *When the Weierstrass modul $f(w)$ is analytically continuable over the circles (C^η), the partial derivatives D_{η_i} are given by* (6.10), (6.14) *and* (6.15) *according as η_i is respectively the radius r_k of C^η_k, the coördinate a_k, or b_k, of the center of C^η_k.*

THEOREM 6.2. *When D is finite at (g, η), a set of necessary and sufficient conditions that the first variation V of D at (g, η) be "identically null" is as follows:*

(i) *For each k, $I[(w - w_k)^2 f(w)]$ tend uniformly to* 0 *as the point on B tends to C^η_k $(k = 0, 1, \cdots, n)$.*

$$\text{(ii)} \qquad \int_{C_k} f(w)\, dw = 0 \qquad (k = 1, \cdots, n),$$

$$\text{(iii)} \qquad \int_{C_k} wf(w)\, dw = 0 \qquad (k = 1, \cdots, n).$$

We assume that $V \equiv 0$. Then condition (i) holds as stated in Theorem 6.1. It follows that $f(w)$ is analytically continuable over C_k so that (6.10), (6.14) and (6.15) hold in accordance with Lemma 6.2. We are assuming that these partial derivatives are all null, so that from (6.14) and (6.15) we infer that

$$\int_{C_k} (w - w_k)f(w)\, ds = 0 \qquad (k = 1, \cdots, n). \tag{6.16}$$

Condition (ii) follows. From the vanishing of the partial derivatives (6.10) and from condition (i) we infer that

$$\int_{C_k} (w - w_k)^2 f(w)\, ds = 0 \qquad (k = 1, \cdots, n). \tag{6.17}$$

Relation (iii) follows and the conditions of the theorem are necessary.

Conversely, we assume that conditions (i), (ii) and (iii) hold. Then $D_{e_k}(g, \eta, \lambda, 0) \equiv 0$ in (λ) by virtue of Theorem 6.1. From (i) it also follows that $f(w)$

is analytically continuable over C_k so that (6.10), (6.14) and (6.15) hold in accordance with Lemma 6.2. Relation (6.16) follows from (ii), and (6.17) from (ii) and (iii). Relations (6.10), (6.14) and (6.15) then show that the derivatives D_{η_i} vanish as stated. Thus the conditions of the theorem are proved sufficient.

The proof of the following lemma is due to Hans Lewy.[21]

LEMMA 6.3. *Under conditions* (i), (ii) *and* (iii) *of Theorem* 6.2, $f(w) \equiv 0$.

Lemma 6.3 and Theorem 6.2 combine in the fundamental theorem.

THEOREM 6.3. *A necessary and sufficient condition that the first variation of a finite Dirichlet sum D be identically zero at* (g, η) *is that the harmonic surface determined by* (g) *on the region with circle parameters* (η) *be minimal.*

The case of two contours, $n = 1$. The fact that two non-intersecting circles can be carried by a conformal map of the plane into two concentric circles suggests that in case $n = 1$, Theorem 6.3 can be replaced by a simpler theorem with a simpler proof. One begins by establishing Theorem 6.1 and equation (6.10) as before, (6.10) holding when the Weierstrass modul $f(w)$ is analytically continuable over the boundary of B. No further calculations are needed.

Take C_0 as the unit circle with center at the origin. Take C_1 as a circle of radius $\sigma < 1$, concentric with C_0. The region B is a ring bounded by C_0 and C_1. Of the circle parameters $(\eta) = (\sigma, u_1, v_1)$, σ alone shall vary. With this understood, set

$$D(g, \eta, \lambda, e) = d(g, \sigma, \lambda, e).$$

Theorem 6.3 is replaced by the following theorem.

THEOREM 6.4. *The harmonic surface determined by* (g) *on B is minimal if, for each admissible* (λ),

$$d_{e_0}(g, \sigma, \lambda, 0) = d_{e_1}(g, \sigma, \lambda, 0) = 0, \tag{6.18}$$

$$d_{\sigma}(g, \sigma, \lambda, 0) = 0. \tag{6.19}$$

As previously, let $f(w)$ be the Weierstrass modul of the harmonic function $h(u, v)$ determined by (g) on B. If (6.18) holds, $I[w^2 f(w)]$ admits a harmonic continuation over C_0 and C_1, and is null on C_0 and C_1 in accordance with Theorem 6.1. Hence $w^2 f(w)$ equals a real constant K on B. If (6.19) also holds, it follows from (6.10) that $K = 0$. Hence $f(w) \equiv 0$, and S is minimal.

THE INSTITUTE FOR ADVANCED STUDY.

[21] See Courant, loc. cit., p. 715. One notes that the conditions of Theorem 6.2 imply that (ii) and (iii) hold even when $k = 0$. This follows from Cauchy's theorem upon integrating $f(w)$ or $wf(w)$ around the boundary of B.

THE LAW OF LARGE NUMBERS FOR CONTINUOUS STOCHASTIC PROCESSES

BY J. L. DOOB

The purpose of this paper is to discuss the law of large numbers as applied to stochastic processes depending on a continuous parameter. In order to treat the temporally homogeneous processes, which will be discussed first, a form of the ergodic theorem of Birkhoff is proved which seems the from best suited to probability applications. In studying differential processes, some theorems on infinite series whose terms are independent chance variables will be needed. These are new and have some independent interest.

1. The ergodic theorem and temporally homogeneous processes. Let X be an abstract space, let $\mathfrak{F}_x$ be a Borel field of X-sets (including the space itself) and let $M(W)$ be a non-negative measure function defined on this Borel field. If $M(X) = +\infty$, we suppose that X is the sum of denumerably many sets of finite measure. Measurability of a numerically-valued function $f(x)$, defined on X, and integration are then defined in the usual way.

Let $f_0(x), f_1(x), \cdots$ be a sequence of measurable functions. Let n be any positive integer; let $E_1, \cdots, E_n$ be any Borel sets of numbers; let $h, \alpha_1, \cdots, \alpha_n$ be integers with $h \geqq 0$ and $0 \leqq \alpha_1 < \cdots < \alpha_n$; let W_h be the X-set determined by the conditions

$$f_{\alpha_j+h}(x) \in E_j \qquad (j = 1, \cdots, n). \tag{1}$$

Then if $M(W_h)$ is independent of h, the sequence $\{f_n(x)\}$ will be said to have the property H.

Let $f_t(x)$ $(0 \leqq t < \infty)$ be a one-parameter family of measurable functions. This family will be said to have the property H if the following condition is satisfied. Let $n, E_1, \cdots, E_n, h, \{\alpha_j\}, W_h$ be as above, except that $h, \alpha_1, \cdots, \alpha_n$ need not be integers. Then $M(W_h)$ is to be independent of h. If in either of the above two definitions, f_j (or f_t) is defined for $j = 0, \pm 1, \pm 2, \cdots$ (or $-\infty < t < \infty$) the definition of the property H remains as above, except that h need no longer be $\geqq 0$. It is evident that the integrability of any $f_j(x)$ in the above implies that of any other, and the value of the integral will be independent of the subscript.

The ergodic theorem of Birkhoff[1] is a theorem about a sequence of functions $f(x), f(Tx), f(T^2x), \cdots$, where $f(x)$ is x-measurable, and T is a measure-preserving point transformation. Such a sequence of functions has the property H.

Received January 14, 1939.

[1] A detailed discussion of the ergodic theorem and related theorems has been given by E. Hopf in his *Ergodentheorie*, Ergebnisse der Mathematik, vol. 5(1937), no. 2.

The following theorem then includes the corresponding form of the ergodic theorem.

THEOREM 1. *Let* $\{f_n(x)\}$ *be a sequence of measurable X-functions with the property* H, *and let* $f_0(x)$ *be integrable. Then* $\lim_{N\to\infty} \frac{1}{N}\sum_1^N f_n(x)$ *exists almost everywhere on* X.

If, in the sequence $\{f_n(x)\}$, n ranges through all integers, Theorem 1 can be proved by finding a measure-preserving set transformation T of the X-sets such that, in a sense easily made precise, $f_j(Tx) = f_{j+1}(x)$ $(j = 0, \pm 1, \cdots)$ and the proof of the ergodic theorem needs no change to be applicable.[2] It is just the other case, however, in which $n \geqq 0$, which is the one most frequently met in probability.

Let Ω be the space of all sequences ω: $(\cdots, \xi_{-1}, \xi_0, \xi_1, \cdots)$ of real numbers. Let m be any positive integer, let $\alpha_1, \cdots, \alpha_m$ be integers, with $\alpha_1 < \cdots < \alpha_m$, and let $E_1, \cdots, E_m$ be Borel sets of numbers. The conditions

$$\xi_{\alpha_j} \in E_j \qquad (j = 1, \cdots, m) \tag{2}$$

determine an Ω-set. Let $\mathfrak{F}_\omega$ be the Borel field of Ω-sets determined by such Ω-sets, and let $\mathfrak{F}_\omega(\alpha_1, \cdots, \alpha_m)$ be the Borel field of Ω-sets determined by such Ω-sets for given $\alpha_1, \cdots, \alpha_m$. We shall define a measure function on the sets of $\mathfrak{F}_\omega$. In order to do this we first define a measure on the sets of $\mathfrak{F}_\omega(\alpha_1, \cdots, \alpha_m)$ for given $\alpha_1, \cdots, \alpha_m$. Choose h so that $\alpha_1 + h \geqq 0$. Let U be the transformation (depending on $\alpha_1, \cdots, \alpha_m, h$), taking the point x of X into every point ω: $(\cdots, \xi_{-1}, \xi_0, \xi_1, \cdots)$ for which $\xi_{\alpha_j} = f_{\alpha_j+h}(x)$. The Ω-set Λ defined by conditions (2) corresponds through the transformation U to an X-measurable set $U^{-1}\Lambda$: that defined by the conditions

$$f_{\alpha_j+h}(x) \in E_j \qquad (j = 1, \cdots, m). \tag{3}$$

Then each set Λ of $\mathfrak{F}_\omega(\alpha_1, \cdots, \alpha_m)$ corresponds under U to a measurable X-set $U^{-1}\Lambda$, and we define $P(\Lambda) = M(U^{-1}\Lambda)$. Defined in this way, $P(\Lambda)$ is a non-negative completely additive function of the sets of $\mathfrak{F}_\omega(\alpha_1, \cdots, \alpha_m)$, with $P(\Omega) \leqq \infty$. Let Λ_N $(N \geqq 1)$ be defined by the conditions

$$\max_{j\leqq m} |\xi_{\alpha_j}| \geqq \frac{1}{N}. \tag{4}$$

Then since each $f_k(x)$ is integrable, $P(\Lambda_N) < \infty$, and if Λ_0 is the Ω-set defined by the conditions $\xi_{\alpha_j} = 0$ $(j = 1, \cdots, m)$, we have $\Omega = \sum_0^\infty \Lambda_n$. The space Ω

[2] Cf. N. Wiener, American Journal of Mathematics, vol. 60(1938), pp. 901–902. The ergodic theorem is usually stated under the hypothesis that the given space has finite measure, but is true in the other case also. Khintchine's proof (Mathematische Annalen, vol. 107(1933), pp. 485–488) holds in the case of infinite measure, with a slight modification at the end.

is thus the sum of denumerably many sets of finite measure, and of a set Λ_0 which may have infinite measure, but no subset of which is in the field $\mathfrak{F}_\omega(\alpha_1, \cdots, \alpha_m)$ of definition. Evidently the definition of $P(\Lambda)$ is independent of the integer h used in defining the transformation U. It is readily verified that if Λ occurs both in the field $\mathfrak{F}_\omega(\alpha_1, \cdots, \alpha_m)$ and in $\mathfrak{F}_\omega(\beta_1, \cdots, \beta_n)$, the P-measure assigned to Λ will be the same in each representation. In particular, Ω appears in each representation (and $P(\Omega) = M(X)$). Then by a theorem of Kolmogoroff[3] (proved under the assumption that $P(\Omega) = 1$, but the proof is valid here) the definition of $P(\Lambda)$ can be extended to all the sets of $\mathfrak{F}_\omega$, giving a completely additive non-negative set function. Consider the Ω-sets determined by the conditions (4), for all $N \geqq 1$, $m \geqq 1$, and all sets of integers $\alpha_1 < \cdots < \alpha_m$. These Ω-sets form a denumerable collection of sets, say M_1, $M_2, \cdots$ when enumerated in some order, each of finite P-measure. Let M_0 be the set containing the single point $(\cdots, 0, 0, 0, \cdots)$. Then $\Omega = \sum_0^\infty M_j$ and $\Omega - M_0$ is the sum of denumerably many sets of finite measure. The theory of measurable functions on Ω goes over in the usual way. A function $f(x)$ which is integrable on Ω must vanish on M_0 if $P(M_0) = +\infty$. If T_ω is the transformation defined by the equations $\xi_j' = \xi_{j+1}$ $(j = 0, \pm 1, \cdots)$, T takes measurable Ω-sets into measurable Ω-sets of the same measure. If $\varphi(\xi_{\alpha_1}, \cdots, \xi_{\alpha_m})$ is a measurable Ω-function, depending only on $\xi_{\alpha_1}, \cdots, \xi_{\alpha_m}$, φ has the same distribution as the measurable X-function $\varphi[f_{\alpha_1+h}(x), \cdots, f_{\alpha_m+h}(x)]$, if h is chosen so that $\alpha_j + h \geqq 0$ $(j = 1, \cdots, m)$, i.e.,

$$(5) \qquad P\{\varphi(\xi_{\alpha_1}, \cdots, \xi_{\alpha_m}) < k\} \equiv M\{\varphi[f_{\alpha_1+h}(x), \cdots, f_{\alpha_m+h}(x)] < k\}.^4$$

In particular, the measurable Ω-function $\xi_j = \xi_j(\omega)$ which takes on the value ξ_j at the point ω: $(\cdots, \xi_{-1}, \xi_0, \xi_1, \cdots)$ has the same distribution as $f_j(x)$, so that $\xi_j(\omega)$ is integrable on Ω since $f_j(x)$, with $f_0(x)$, is integrable on X. Then by the ergodic theorem of Birkhoff,

$$(6) \qquad \lim_{n\to\infty} \frac{1}{n} \sum_0^{n-1} \xi_1(T^j\omega) = \lim_{n\to\infty} \frac{1}{n} \sum_1^n \xi_j(\omega)$$

exists almost everywhere on Ω, or, in terms of X-measure,

$$(6') \qquad \lim_{n\to\infty} \frac{1}{n} \sum_1^n f_j(x)$$

exists almost everywhere on X, as was to be proved.

In the following theorem we shall be discussing a function $f(t, x)$ of the real variable t and the variable x in X-space. Measure in (t, x)-space is supposed to have been defined in the usual way, so that the (t, x)-measure of the direct

[3] *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Ergebnisse der Mathematik, vol. 2(1933), no. 3, pp. 27–30.

[4] Throughout this paper, if a point set is determined by some set of conditions C, the set will be denoted by $\{C\}$.

product of a Lebesgue measurable t-set with a measurable X-set is the product of their measures.

THEOREM 2. *Let $f(t, x)$ be a measurable (t, x)-function which is X-measurable for each fixed value of $t \geqq 0$ and such that the one-parameter family of X-functions $\{f(t, x)\}$ $(0 \leqq t < \infty)$ has the property* H. *Then if $f(0, x)$ is X-integrable,*

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f(t, x)\, dt$$

exists almost everywhere on X.

This theorem is reduced to Theorem 1 just as in the usual proof of the ergodic theorem.[5]

Let $\{x_t\}$ be a one-parameter family of chance variables, $-\infty < t < \infty$.[6] It is supposed that if $\alpha_1, \cdots, \alpha_n$ are any distinct values of t, and if $\lambda_1, \cdots, \lambda_n$ are real numbers, the probability of the simultaneous truth of the inequalities

$$x_{\alpha_j} < \lambda_j \qquad (j = 1, \cdots, n) \tag{7}$$

is given. We shall frequently suppose that this probability is independent of translations of the t-axis, i.e., that the stochastic process is temporally homogeneous. Since the chance variable x_t is not given a priori as a function of t and a probability parameter, temporal homogeneity and the existence of[7] Ex_0 do not of themselves imply, if we use the ergodic theorem, that $\lim_{T\to\infty} T^{-1}\int_0^T x_t\, dt$ exists with probability 1.[8] The justification and meaning of this application of the ergodic theorem will now be discussed. The author has discussed in detail the reduction of probability relations on a continuous stochastic process to measure relations on an abstract space.[9] The appropriate space is a suitably chosen subspace Ω of the space Ω^* of real-valued functions $x(t)$, defined for all t. The Ω-measure is determined by assigning as the measure of the Ω-set

$$\{x(\alpha_j) < \lambda_j \quad (j = 1, \cdots, n)\} \tag{8}$$

(i.e., the set of functions $x(t)$ determined by the conditions between the braces) the probability that (7) is true. If almost all functions $x(t)$ have some property, it is said to be true of the x_t, with probability 1. Measurability of a process is defined as follows. Let $f[\tau, x(t)]$ be the function, defined on the product space $T \times \Omega$ of the t-axis with Ω, which takes on the value $x_0(t_0)$ at the

[5] E. Hopf, ibid., pp. 53–54.

[6] We allow t to range through all values, to simplify the notation somewhat, but the following discussion will remain valid if t is restricted to be non-negative.

[7] Throughout this paper, if x is a chance variable, Ex will denote its expectation.

[8] Cf., however, Wiener, American Journal of Mathematics, vol. 60(1938), pp. 899–901, who, in discussing a similar problem from another point of view, supposes just the explicit definition of x_t as a measurable function of two variables (t and a probability parameter) which we do not.

[9] Transactions of the American Mathematical Society, vol. 42(1937), pp. 107–140.

point $[t_0, x_0(t)]$ of $T \times \Omega$, so that we shall sometimes write $f = x(\tau)$. Then the process is called measurable if, when measure is defined on $T \times \Omega$ in the usual way, as the product of Lebesgue measure on t-space and P-measure on Ω, $x(\tau)$ becomes a measurable $T \times \Omega$-function. In the temporally homogeneous case, there is a space $\Omega \subset \Omega^*$ which is that of a measurable process, if and only if

$$\lim_{h\to 0} P\{|x_{t+h} - x_t| > \epsilon\} = 0 \tag{9}$$

for all t, and every $\epsilon > 0$.[10] In the general case, according to a letter from Kolmogoroff, a necessary and sufficient condition is that, except for an exceptional t-set of measure 0, (9) holds, for every $\epsilon > 0$, when $h \to 0$ on an h-set (which may depend on t but not on ϵ) having metric density 1 at $h = 0$.[11] If the process with space Ω is measurable, almost all functions $x(t)$ in Ω are Lebesgue measurable,[12] i.e., in the usual terminology, x_t is Lebesgue measurable in t, with probability 1.

THEOREM 3. *If the chance variables $\{x_t\}$ are those of a measurable temporally homogeneous process, and if Ex_0 exists, then $\lim_{T\to\infty} T^{-1}\int_0^T x_t\,dt$ exists, with probability 1, i.e., $\lim_{T\to\infty} T^{-1}\int_0^T x(t)\,dt$ exists for almost all functions $x(t)$ on the space Ω of the measurable process.*

The first statement of the theorem is in the usual convenient and suggestive terminology, but has no meaning a priori, because $\int_0^T x_t\,dt$ is an undefined symbol. As appears below, $\int_0^T x(t)\,dt$ is a P-measurable Ω-function, i.e., a chance variable, and $\int_0^T x_t\,dt$ is to be interpreted as this chance variable. It may also be possible, in some cases, to define $\int_0^T x_t\,dt$ in other ways, in which the integral sign is no longer interpreted literally.[13] As formulated here, the proof is an immediate corollary of Theorem 2. The function $x(\tau) = f[\tau, x(t)]$ is a measurable function of the two variables τ, $x(t)$, which, as τ varies, generates a one-parameter family of functions with the property H, and the function $f[0, x(t)]$ is integrable on Ω: Ex_0 exists.

2. **Differential processes.** A stochastic process determined by the chance variables $\{x_t\}$ is called a differential process if whenever $t_1 < t_2 < \cdots < t_\nu$

[10] Doob, ibid., p. 117.

[11] Cf. also Warren Ambrose, Transactions of the American Mathematical Society, vol. 47(1940), p. 71.

[12] Doob, Transactions of the American Mathematical Society, vol. 42(1937), p. 113.

[13] Doeblin, Thesis, University of Paris, 1938, pp. 109–115; Cramér, Annals of Mathematics, vol. 41(1940), p. 219.

$(\nu > 1)$, $x_{t_2} - x_{t_1}, \cdots, x_{t_\nu} - x_{t_{\nu-1}}$ form an independent set of chance variables. In this case, we shall call the process temporally homogeneous if the probability of an inequality of the form $x_{s_2} - x_{s_1} < k$ is independent of translations of the t-axis. The only relations of interest are relations between differences $x_s - x_t$ and it is unnecessary even to define probabilities of relations such as $x_t < k$. If such probabilities are not defined, x_t is not a chance variable at all, and the chance variables which determine the probability relations are $\{x_t - x_0\}$. However, it is usually convenient to define the probabilities of events such as (7) above. This can be done by giving x_0 some distribution, and prescribing that the chance variable x_0 thus defined is to be independent of every chance variable $x_s - x_t$ with $0 \leqq t < s$ or $0 \geqq t > s$, or any P-measurable function of such chance variables. For instance, we can set $x_0 = 0$ with probability 1. Since the conditions for measurability of a process depend only on differences $x_s - x_t$, the specific distribution given to x_0 will be irrelevant to measurability.

By a theorem of Lévy,[14] if a process is differential, there is a function $\alpha(t)$ such that if the chance variable $y_t = x_t - \alpha(t)$ is substituted for x_t, then if t is not in some exceptional denumerable set (which will be denoted by S_0 throughout the remainder of this paper) $t_n \to t$ implies that $y(t_n) \to y(t)$, with probability 1, and even if $t \in S_0$, the limit exists, for monotone approach:

$$\lim_{t_n \uparrow t} y_{t_n} = y_{t-0}, \qquad \lim_{t_n \downarrow t} y_{t_n} = y_{t+0} \tag{10}$$

with probability 1. The limits y_{t+0}, y_{t-0} defined by (10) are independent of the sequences $\{t_n\}$, neglecting 0 probabilities. Moreover if $t \in S_0$, neither $y_{t+0} - y_t$ nor $y_t - y_{t-0}$ is constant $(\neq 0)$ with probability 1. A differential process for which we can take $\alpha(t) \equiv 0$ is called centered. The set S_0 is independent of the centering function $\alpha(t)$, which is not uniquely determined. Several results will now be proved about sums of independent chance variables, which will be useful in studying differential processes.

If $E(x_t - x_0)$ exists for some value t_0 of t, the following lemma shows that $E(x_t - x_0)$ exists for every value of t between 0 and t_0.

LEMMA 1. *If x, y are independent chance variables, and if $E(x + y)$ exists, then Ex, Ey exist also.*

If x, y have distribution functions $F(x)$, $G(y)$, $E(x + y)$ becomes

$$\int\int (x + y)\, dF(x)\, dG(y)$$

which is by hypothesis an absolutely convergent integral. The double integral can then be evaluated in terms of the iterated integrals. The evaluation implies the existence of Ex, Ey, and gives the formula $E(x + y) = Ex + Ey$.

[14] Annali della Scuola Normale Superiore di Pisa, (2), vol. 3(1934), pp. 342–343. Cf. also Doob, Transactions of the American Mathematical Society, vol. 42(1937), pp. 126–129.

LEMMA 2. *If x, y are independent chance variables, whose expectations exist, and if $Ey = 0$, then $E\,|\,x\,| \leqq E\,|\,x + y\,|$.*

Let E_x denote expectation for x given. Then since $x = (x + y) - y$,

$$E_x x = x = E_x(x + y) - E_x y = E_x(x + y) - Ey = E_x(x + y), \tag{11}$$

so that

$$|\,x\,| \leqq E_x\,|\,x + y\,|. \tag{12}$$

Taking expectations on both sides, we have

$$E\,|\,x\,| \leqq E\,|\,x + y\,|, \tag{13}$$

as was to be proved.

LEMMA 3. *Let $x_1, \cdots, x_n$ be an independent set of chance variables. Suppose that x_n, $x_n + x_{n-1}, \cdots, x_n + \cdots + x_1$ have median 0. Let $Q_n(l)$ be the function of concentration of $s_n = x_1 + \cdots + x_n$:*

$$Q_n(l) = \underset{-\infty<a<\infty}{\text{L.U.B.}}\, P\{a \leqq s_n \leqq a + l\}. \tag{14}$$

Then[15]

$$P\{\max_{j \leqq n} |\,x_1 + \cdots + x_j\,| > 2l\} \leqq 2[1 - Q_n(l)]. \tag{15}$$

LEMMA 4. *Let $x_1, x_2, \cdots$ be mutually independent chance variables, and suppose that $\sum_1^\infty x_j = x$ with probability 1. Let $x_n + \cdots + x_{n'}$ have median 0 for all $n' > n$. Then if $Q(l)$ is the function of concentration of x,*

$$P\{\underset{n \geqq 1}{\text{L.U.B.}}\, |\,x_1 + \cdots + x_n\,| > 2l\} \leqq 2\{1 - Q(l)\}. \tag{16}$$

If Ex exists, then $E(\underset{n \geqq 1}{\text{L.U.B.}}\, |\,x_1 + \cdots + x_n\,|)$ also exists.

Let $Q_n(l)$ be the function of concentration of $x_1 + \cdots + x_n$. Since adding independent chance variables cannot increase the function of concentration,[16] $Q_n(l) \geqq Q(l)$. Then the inequality (15) above is true, and implies (16), when $n \to \infty$. Now suppose that Ex exists. Consider the following sum:

$$\begin{aligned} \sum_{n=1}^\infty P\{\underset{j \geqq 1}{\text{L.U.B.}}\, |\,x_1 + \cdots + x_j\,| > n\} &\leqq 2 \sum_{n=1}^\infty [1 - Q(\tfrac{1}{2}n)] \\ &\leqq 2 \sum_{n=1}^\infty [1 - P\{|\,x\,| \leqq \tfrac{1}{4}n\}] \\ &\leqq 2 \sum_{n=1}^\infty P\{4\,|\,x\,| > n\}. \end{aligned} \tag{17}$$

[15] Paul Lévy, *Théorie de l'Addition des Variables Aléatoires*, Paris, 1937, pp. 137–138. (Lévy's notation is somewhat different from ours.)

[16] Lévy, ibid., pp. 89–90.

If z is any chance variable, Ez exists if and only if $\sum_1^\infty P\{|z| > n\} < \infty$.[17] Then the above sums converge (since $E(4x)$ exists), so that $E(\text{L.U.B.}_{j \geqq 1} |x_1 + \cdots + x_j|)$ exists.

LEMMA 5. *Let $x_1, x_2, \cdots$ be a sequence of chance variables, and let x be any chance variable dependent upon them. Suppose that Ex exists, and let $E_n x$ (a function of $x_1, \cdots, x_n$) be the conditional expectation of x for given $x_1, \cdots, x_n$. Then $\lim_{n \to \infty} E_n x = x$ with probability* 1.

If x takes on only the values 0, 1, so that it can be considered to characterize an event, the event $x = 1$, this lemma becomes a well-known result due to Lévy, and a suitable modification of Lévy's proof[18] gives the general case.

THEOREM 4. *Let $x_1, x_2, \cdots$ be independent chance variables, and let $\sum_1^\infty x_j = x$, with probability* 1. *Then if Ex exists, it follows that Ex_n exists for all n, that*

$$Ex = \sum_1^\infty Ex_n \tag{18}$$

and that $E(\text{L.U.B.}_{j \geqq 1} |x_1 + \cdots + x_j|)$ exists.[19]

This theorem provides a converse to Lebesgue's theorem on term by term integration of a convergent series of functions. According to Lebesgue's theorem, if $f_1, f_2, \cdots$ are measurable functions, if $\sum_1^\infty f_j = s$ almost everywhere, and if the partial sums $\{s_n\}$ are bounded by an integrable function $\varphi \geqq 0$: $|s_n| \leqq \varphi$, $n \geqq 1$, then the series can be integrated term by term. Lebesgue's condition, the existence of φ, is not necessary, in general, but Theorem 4 shows that if the functions $f_1, f_2, \cdots$ are independent, then Lebesgue's condition is necessary. In fact, if the sum s is integrable, then the partial sums are bounded by the integrable function $\text{L.U.B.}_{j \geqq 1} |s_j|$.

Proof of Theorem 4. Since $x = x_n + \sum_{j \neq n} x_j$, where x_n and the sum are independent, it follows from Lemma 1 that Ex_n exists. Using the notation of Lemma 4, we have

$$\begin{aligned} E_n x &= E_n(x_1 + \cdots + x_n) + E_n(x_{n+1} + \cdots) \\ &= x_1 + \cdots + x_n + E(x_{n+1} + \cdots). \end{aligned} \tag{19}$$

[17] In fact it is easily verified that

$$\sum_1^\infty P\{|z| > n\} \leqq E|z| \leqq \sum_0^\infty P\{|z| > n\}.$$

[18] Ibid., p. 129.

[19] This theorem, with $Ex_1 = Ex_2 = \cdots = 0$, is contained in a theorem of J. Marcinkiewicz and A. Zygmund, Studia Mathematica, vol. 7(1938), p. 113.

Then

$$E_n x - (x_1 + \cdots + x_n) = Ex - (Ex_1 + \cdots + Ex_n), \tag{20}$$

so that, if we let $n \to \infty$, the left approaches 0, according to Lemma 5, with probability 1. Thus $Ex - \sum_1^n Ex_j \to 0$, and (18) is proved. Moreover, from (20),

$$\begin{aligned} |x_1 + \cdots + x_n| &\leqq |E_n x| + |Ex - (Ex_1 + \cdots + Ex_n)| \\ &\leqq E_n |x| + |Ex - (Ex_1 + \cdots + Ex_n)|, \end{aligned} \tag{21}$$

so that

$$E|x_1 + \cdots + x_n| \leqq E|x| + |Ex - (Ex_1 + \cdots + Ex_n)| \leqq K, \qquad n \geqq 1, \tag{22}$$

for suitably chosen K. Now let $\xi_1, \xi_2, \cdots$ be chance variables independent of each other, and of the x's, and such that for each j, ξ_j has the same distribution as x_j. Let $\xi = \sum_1^\infty \xi_j$, and let $y = x_j - \xi_j$. Then the chance variables $y_1, y_2, \cdots$ are mutually independent and any sum of y's has a symmetric distribution with median 0. By Lemma 4, if $z = \text{L.U.B.}_{j \geqq 1} |y_1 + \cdots + y_j|$, Ez exists. We have

$$|x_1 + \cdots + x_n| \leqq z + |\xi_1 + \cdots + \xi_n|. \tag{23}$$

If we take the conditional expectation of both sides for $x_1, x_2, \cdots$ fixed, we find (if w is the conditional expectation of z for $x_1, x_2, \cdots$ fixed),

$$|x_1 + \cdots + x_n| \leqq w + E|\xi_1 + \cdots + \xi_n| \leqq w + K. \tag{24}$$

Now Ew exists ($Ew = Ez$), so that $E[\text{L.U.B.}_{j \geqq 1} |x_1 + \cdots + x_j|] \leqq Ew + K < \infty$, as was to be proved.

THEOREM 5. *Let $\{x_t\}$ be the chance variables of a differential process, and suppose that $E(x_t - x_0)$ exists for arbitrarily large values of t and $-t$. Then $E(x_t - x_0) = e(t)$ exists for all t, and the process with variables $\{y_t = x_t - x_0 - e(t)\}$ is centered.*

The fact that $e(t)$ exists for all t is a trivial consequence of Lemma 1. In proving the second part of the theorem we can suppose that $e(t) \equiv 0$ (replacing x_t by $x_t - x_0 - e(t)$ unless $e(t) \equiv 0$ already) and prove that the process is centered. If $\alpha(t)$ is a function which is known to center the process (so that the process with variables $\{x_t - \alpha(t)\}$ is centered), we shall show that (a) $\alpha(t)$ is continuous if $t \,\epsilon'\, S_0$; and (b) if $t \to t'$ monotonely, $\lim_{t \to t'} \alpha(t)$ exists (t' arbitrary). Since (a) and (b) imply that if $t \,\epsilon'\, S_0$, $\lim_{t_n \to t} x_{t_n} = x_t$ with probability 1, and that for any t, if $t_n \to t$ monotonely, $\lim_{n \to \infty} x_{t_n}$ exists with probability 1, only a third

fact need be proved: (c) if $t_n \to t$ monotonely, and if $\lim_{n\to\infty} x_{t_n} = x_t + c$ with probability 1 (where c is a constant), then $c = 0$.

Proof of (a). It is sufficient to show that if $t_n \to t \,\epsilon'\, S_0$, and if the sequence $\{t_n\}$ is monotone, then $\alpha(t_n) \to \alpha(t)$. Suppose for definiteness that the sequence $\{t_n\}$ is monotone increasing. Then since $t \,\epsilon'\, S_0$, $x_{t_n} - \alpha(t_n) \to x_t - \alpha(t)$ with probability 1, so that

$$(25) \qquad x_t - x_{t_1} - [\alpha(t) - \alpha(t_1)] = \sum_1^\infty [x_{t_{j+1}} - x_{t_j} - \alpha(t_{j+1}) + \alpha(t_j)]$$

with probability 1. By Theorem 4, since $E[x_t - x_{t_1} - \alpha(t) + \alpha(t_1)] = -\alpha(t) + \alpha(t_1)$ exists, we can sum the expectations in (25) term by term:

$$(26) \qquad -\alpha(t) + \alpha(t_1) = \sum_1^\infty [-\alpha(t_{j+1}) + \alpha(t_j)],$$

i.e., $\alpha(t_n) \to \alpha(t)$, as was to be proved.

Proof of (b). Suppose that $t_n \to t$, and that the sequence $\{t_n\}$ is monotone, say monotone increasing. Since $x_t - x_{t_1} = (x_t - x_{t_n}) + (x_{t_n} - x_{t_1})$ with $E(x_t - x_{t_n}) = 0$, Lemma 2 implies that

$$(27) \qquad E \mid x_{t_n} - x_{t_1} \mid \leqq E \mid x_t - x_{t_1} \mid, \qquad n \geqq 1.$$

Then there is a number K so large that

$$(28) \qquad P\{\mid x_{t_n} - x_{t_1} \mid < K\} > \tfrac{1}{2}, \qquad n \geqq 1.$$

Now $\lim_{n\to\infty} [x_{t_n} - \alpha(t_n)]$ is supposed to exist, with probability 1, so that we can increase K, if necessary, so that (28) will be true, and also

$$(29) \qquad P\{\mid x_{t_n} - x_{t_1} - \alpha(t_n) + \alpha(t_1) \mid < K\} > \tfrac{1}{2}, \qquad n \geqq 1.$$

Equations (28) and (29) imply that

$$(30) \qquad \mid \alpha(t_n) - \alpha(t_1) \mid < 2K, \qquad n \geqq 1.$$

Since $x_t - x_{t_1} - \alpha(t_n) + \alpha(t_1) = (x_t - x_{t_n}) + [x_{t_n} - x_{t_1} - \alpha(t_n) + \alpha(t_1)]$, with $E(x_t - x_{t_n}) = 0$, Lemma 2 implies that

$$(31) \qquad \begin{aligned} E \mid x_{t_n} - x_{t_1} - \alpha(t_n) + \alpha(t_1) \mid &\leqq E \mid x_t - x_{t_1} - \alpha(t) + \alpha(t_1) \mid \\ &\leqq E \mid x_t - x_{t_1} \mid + 2K, \qquad n \geqq 1. \end{aligned}$$

The fact that the left side of (31) is bounded uniformly for all n implies, by a theorem of Fatou,[20] that

$$E \left| \lim_{n\to\infty} [x_{t_n} - x_{t_1} - \alpha(t_n) + \alpha(t_1)] \right| = E \left| \sum_1^\infty [x_{t_{n+1}} - x_{t_n} - \alpha(t_{n+1}) + \alpha(t_n)] \right|$$

exists. Then by Theorem 4,

$$\lim_{n\to\infty} \alpha(t_n) = \alpha(t_1) - \sum_1^\infty E[x_{t_{n+1}} - x_{t_n} - \alpha(t_{n+1}) + \alpha(t_n)]$$

exists, as was to be proved.

[20] E. C. Titchmarsh, *The Theory of Functions*, Oxford, 1932, p. 346.

Proof of (c). Suppose that $t_n \to t$ monotonely, and that $x_{t_n} \to x_t + c$, with probability 1. Suppose for definiteness that the sequence $\{t_n\}$ is monotone increasing. Then

$$(32) \qquad x_{t_n} - x_{t_1} = \sum_1^{n-1} (x_{t_{j+1}} - x_{t_j}) \to x_t - x_{t_1} + c,$$

so, by Theorem 4, $0 = E(x_{t_n} - x_{t_1}) \to E(x_t - x_{t_1} + c) = c$: $c = 0$, as was to be proved.

In the following discussion, we shall frequently use the concept $\underset{a \leq t \leq b}{\text{L.U.B.}} | x_t - x_0 |$, where the variables $\{x_t\}$ are those of a centered differential process. This symbol does not represent a chance variable a priori, i.e., $P\{\underset{a \leq t \leq b}{\text{L.U.B.}} | x_t - x_0 | > k\}$ cannot be defined directly in terms of the given probabilities of the process. There have been two ways suggested for handling this problem. The first method, that of Khintchine,[21] defines $P\{\underset{a \leq t \leq b}{\text{L.U.B.}} | x_t - x_0 | > k\}$ as L.U.B. $P\{\max_j | x_{t_j} - x_0 | > k\}$ for every possible finite set of numbers t_1, $t_2, \cdots$ in the interval $a \leq t \leq b$. Thus Khintchine does not actually define $\underset{a \leq t \leq b}{\text{L.U.B.}} | x_t - x_0 |$ as a chance variable, but the latter can be defined, in the spirit of his definition as follows. If $t_1, t_2, \cdots$ is any sequence of numbers in the interval $a \leq t \leq b$, $\xi = \underset{j}{\text{L.U.B.}} | x_{t_j} - x_0 |$ is a chance variable depending on $t_1, t_2, \cdots$. It is readily shown that there is a choice of $t_1, t_2, \cdots$ maximizing E (arc tan ξ). Such a choice is not unique, but determines ξ uniquely, neglecting 0 probabilities, and $x = \underset{j}{\text{L.U.B.}} | x_{t_j} - x_0 |$ can be defined as one of these maximal ξ's. Then $P\{\underset{a \leq t \leq b}{\text{L.U.B.}} | x_t - x_0 | > k\}$ as defined above becomes simply $P\{x > k\}$. The second method of handling the symbol $\underset{a \leq t \leq b}{\text{L.U.B.}} | x_t - x_0 |$ uses the literal interpretation of L.U.B. Probability relations are reduced to measure relations on a space Ω of functions ω: $x(t)$, as discussed above. For each function $x(t)$, $f(\omega) = \underset{a \leq t \leq b}{\text{L.U.B.}} | x(t) - x(0) |$ has its usual meaning, and all that need be shown is that $f(\omega)$ is a measurable Ω-function. In many cases, as in those below, the functions $x(t)$ in Ω can be assumed to have the property that there is a denumerable sequence of points $\{t_j\}$ in any given t-interval I such that

$$(33) \qquad \underset{t \in I}{\text{L.U.B.}}\, x(t) = \underset{j}{\text{L.U.B.}}\, x(t_j), \qquad \underset{t \in I}{\text{G.L.B.}}\, x(t) = \underset{j}{\text{G.L.B.}}\, x(t_j).$$

If this is true, the space Ω is called quasi-separable. A chance variable is a measurable Ω-function, and evidently if Ω is quasi-separable, since for given $t = t_0$, $x(t_0)$ is a chance variable, $\underset{a \leq t \leq b}{\text{L.U.B.}} | x(t) - x(0) | = \underset{j}{\text{L.U.B.}} | x(t_j) - x(0) |$

[21] Ergebnisse der Mathematik, vol. 2(1933), no. 4, *Asymptotische Gesetze der Wahrscheinlichkeitsrechnung*, pp. 68–69.

is also a chance variable—and this chance variable will be written $\underset{a \leqq t \leqq b}{\text{L.U.B.}} \mid x_t - x_0 \mid$. Evidently the two interpretations of $\underset{a \leqq t \leqq b}{\text{L.U.B.}} \mid x_t - x_0 \mid$ are essentially the same, where the second is applicable, i.e., in the quasi-separable case. It has been shown[22] that if a differential process is centered, a quasi-separable space Ω can always be found for the process, and that the process will then be measurable.

THEOREM 6. *Let $\{x_t\}$ be the chance variables of a centered differential process, and suppose that $E(x_t - x_0)$ exists for arbitrarily large values of t and $-t$. Then if I is any finite t-interval,* $\underset{t \in I}{\text{L.U.B.}} \mid x_t - x_0 \mid$ *is a chance variable whose expectation exists.*

Theorems 5 and 6 provide the complete analogue of Theorem 4, for continuous processes. The proof of Theorem 6 is parallel to that of Theorem 4, so the steps will only be sketched. It will be sufficient to prove the theorem for I an interval having the origin as endpoint—say as left endpoint, so that I is the interval $0 \leqq t \leqq T$. We note first that according to Theorem 5 the function $e(t) = E(x_t - x_0)$ can be used to center the process, so that $e(t)$ is continuous for $t \,\epsilon'\, S_0$ and if $t \to t'$ monotonely, $\lim_{t \to t'} e(t)$ exists, for any t'. Then $e(t)$ must be a bounded function for t in I, $\mid e(t) \mid \leqq K_1$. Since

$$x_T - x_0 - e(T) + e(t) = x_t - x_0 + [x_T - x_t - e(T) + e(t)],$$

it follows from Lemma 2 that

$$\begin{aligned} E \mid x_t - x_0 \mid &\leqq E \mid x_T - x_0 \mid + \mid e(T) - e(t) \mid \qquad (0 \leqq t \leqq T) \\ &\leqq E \mid x_T - x_0 \mid + 2K_1 = K. \end{aligned} \tag{34}$$

Let $\{\xi_t\}$ be the chance variables of a differential process independent of the x_t-process but with the same distribution. Let $y_t = x_t - \xi_t - x_0 + \xi_0$. By an easy extension of Lemma 4 to continuous processes, if $z = \underset{t \in I}{\text{L.U.B.}} \mid y_t \mid$, Ez exists. We have then $\mid x_t - x_0 \mid \leqq z + \mid \xi_t - \xi_0 \mid$ corresponding to (23), and can then go on as in the proof of Theorem 4 to find that $E(\underset{t \in I}{\text{L.U.B.}} \mid x_t - x_0 \mid) \leqq Ez + K$, completing the proof.

In the case of a temporally homogeneous differential process, some of the preceding results are considerably simpler. In this case, if $E(x_1 - x_0)$ is supposed to exist, $E(x_n - x_0) = nE(x_1 - x_0)$ also exists ($n = \pm 1, \pm 2, \cdots$). Then $e(t)$ exists for all t, and in fact $E(x_t - x_s)$ exists for all s, t, and is $e(t) - e(s)$. The function $e(t)$ satisfies the functional equation

$$e(s + t) = e(s) + e(t). \tag{35}$$

[22] Doob, Transactions of the American Mathematical Society, vol. 42(1937), pp. 134–135. In an article to appear in the Transactions of the American Mathematical Society (1940), it is shown that no matter what the given probability relations, there is always a quasi-separable process.

If the process is centered, $e(t)$ is continuous,[23] and the only solution of (35) is $e(t) = at$, where $a = e(1) = E(x_1 - x_0)$. The following theorem shows what a small restriction it is to suppose that a temporally homogeneous process is centered.

THEOREM 7. *A measurable temporally homogeneous differential process is centered.*

Let $\varphi_t(z)$ be the characteristic function of $x_t - x_0$:

$$\varphi_t(z) = E\{e^{iz(x_t - x_0)}\}. \tag{36}$$

Then since the characteristic function of the sum of independent chance variables is the product of their characteristic functions, the equation

$$(x_s - x_0) + (x_{s+t} - x_s) = x_{s+t} - x_0$$

implies that

$$\varphi_s(z) \cdot \varphi_t(z) = \varphi_{s+t}(z) \tag{37}$$

for all z. Since $\varphi_t(0) = 1$, there is a positive number K such that $\varphi_1(z) \neq 0$, if $|z| \leqq K$. Then let s be any positive number, and let ν be a positive integer larger than s. We have $\varphi_\nu(z) = \varphi_1(z)^\nu$, so that $\varphi_\nu(z) \neq 0$, if $|z| \leqq K$, and since $\varphi_\nu(z) = \varphi_s(z) \cdot \varphi_{\nu - s}(z)$, $\varphi_s(z) \neq 0$, if $|z| \leqq K$. Then for each t, the function $\psi_t(z) = \log \varphi_t(z)$ can be defined to be single valued and continuous in z ($|z| \leqq K$), with $\psi_t(0) = 0$. The functional equation becomes

$$\psi_s(z) + \psi_t(z) = \psi_{s+t}(z). \tag{38}$$

Since the process is measurable, $\exp\{iz[x(\tau) - x(0)]\}$ (with $x(\tau) = f[\tau, x(t)]$ as defined above), for fixed z, is a measurable function of $[\tau, x(t)]$ in the product space $T \times \Omega$, so that by Fubini's theorem, $\varphi_t(z)$, and therefore $\psi_t(z)$, is a Lebesgue measurable function of t. The only measurable solutions of (38) are known to be of the form $\psi_t(z) = A(z)t$.[24] Then $\varphi_t(z) = e^{A(z)t}$ for $|z| \leqq K$, where $A(z) = \psi_1(z)$ is continuous for $|z| \leqq K$. Thus the characteristic function of $x_{t+h} - x_t$, $e^{A(z)h}$, converges uniformly for $|z| \leqq K$, when $h \to 0$, to the function 1. Then $x_{t+h} - x_t \to 0$ in probability, when $h \to 0$,[25] and this evidently implies that the process is centered.

[23] Evidently S_0 is empty if the given differential process is centered and temporally homogeneous. The continuity of $e(t)$ then follows from the fact that $e(t)$ is a centering function, or the continuity can be deduced directly if we use Theorem 4.

[24] W. Sierpinski, Fundamenta Mathematicae, vol. 1(1920), pp. 116–122.

[25] Lévy, *Théorie de l'Addition des Variables Aléatoires*, Paris, 1937, pp. 48–50, shows that if a sequence of chance variables has the property that the corresponding characteristic functions converge, and converge uniformly in some neighborhood of $z = 0$, then the distribution functions of the given chance variables converge to a distribution function with characteristic function the limit of the given characteristic function. His proof applies to the present case, in which the characteristic functions $\{\varphi_h(z)\}$ are not known to converge for all z (when $h \to 0$), but in which the limit function is known to be identically 1 near $z = 0$. The limiting distribution function must be that of a chance variable which is 0 with probability 1. The asserted convergence in probability follows from this.

LEMMA 6. *Let $\{u_t\}$, $\{v_t\}$ be chance variables defined for t in some set E unbounded on the right. Suppose that the $\{u_t\}$ are independent of the $\{v_t\}$. Then*

(i) *if $u_t - v_t \to 0$ $(t \to \infty)$ in probability, there is a numerically-valued function e_t such that $u_t - e_t \to 0$, $v_t - e_t \to 0$ $(t \to \infty)$ in probability;*

(ii) *if E is denumerable, then* (i) *holds when "convergence in probability" is replaced in hypothesis and conclusion by "convergence with probability 1".*

Proof of (i). Let $Q_t(l)$ be the function of concentration of u_t, and let $\bar{Q}_t(l)$ be that of $u_t - v_t$. Since u_t, v_t are independent, $Q_t(l) \geqq \bar{Q}_t(l)$.[26] Since $u_t - v_t \to 0$ in probability, $\bar{Q}_t(l) \to 1$, for all $l > 0$. Then $Q_t(l) \to 1$ for all $l > 0$, so that if e_t is a median of u_t, and if $\epsilon > 0$, $P\{| u_t - e_t | < \epsilon\} \to 1$. Thus $u_t - e_t \to 0$ in probability, and therefore $v_t - e_t = -(u_t - v_t) + (u_t - e_t) \to 0$ in probability also.

Proof of (ii). If $u_t - v_t \to 0$ with probability 1, $u_t - v_t \to 0$ in probability, so by (i), $u_t - e_t \to 0$, $v_t - e_t \to 0$, in probability, if e_t is suitably chosen. Let ϵ be a positive number, and let $\bar{t}$ be so large that if $t > \bar{t}$, $P\{| v_t - e_t | < \frac{1}{2}\epsilon\} > \frac{1}{2}$. Let $t_1, t_2, \cdots$ be the values of $t > \bar{t}$ in E. Then

$$
\begin{aligned}
P\{\underset{t>\bar{t}}{\mathrm{L.U.B.}}\, | u_t - v_t | > \tfrac{1}{2}\epsilon\}
&\geqq \sum_{k=1}^{\infty} P\{| u_{t_j} - e_{t_j} | \leqq \epsilon, j < k; | u_{t_k} - e_{t_k} | > \epsilon; | v_{t_k} - e_{t_k} | < \tfrac{1}{2}\epsilon\} \\
&= \sum_{k=1}^{\infty} P\{| u_{t_j} - e_{t_j} | \leqq \epsilon, j < k; | u_{t_k} - e_{t_k} | > \epsilon\} \cdot P\{| v_{t_k} - e_{t_k} | < \tfrac{1}{2}\epsilon\} \qquad (39) \\
&> \tfrac{1}{2} \sum_{k=1}^{\infty} P\{| u_{t_j} - e_{t_j} | \leqq \epsilon, j < k; | u_{t_k} - e_{t_k} | > \epsilon\} \\
&= \tfrac{1}{2} P\{\underset{t>\bar{t}}{\mathrm{L.U.B.}}\, | u_t - e_t | > \epsilon\}.
\end{aligned}
$$

The quantity on the left approaches 0 with $1/t$ by hypothesis, so that on the right also approaches 0, i.e., $u_t - e_t \to 0$ (and also $v_t - e_t = -(u_t - v_t) + (u_t - e_t) \to 0$) with probability 1, as was to be proved.

THEOREM 8. *Let $\{x_t\}$ be the variables of a centered differential process. Then there is a numerically-valued function $f(t)$ such that*

$$\lim_{t \to \infty} \left[\frac{x_t - x_0}{t} - f(t) \right] = 0 \qquad (40)$$

with probability 1, *if and only if there is a sequence of constants $\{f_n\}$ such that*

$$\lim_{n \to \infty} \left[\frac{x_n - x_0}{n} - f_n \right] = 0 \qquad (41)$$

with probability 1.

[26] Lévy, ibid., pp. 89–90.

Evidently if (40) is true, (41) is also, with $f_n = f(n)$. Conversely, suppose that (41) is true. Let $\{\xi_t\}$ be chance variables of a stochastic process independent of the x_t-process, but with exactly the same probability distribution. Then if $y_t = x_t - \xi_t$, the y_t determine a differential process. Evidently (41) implies that

$$\lim_{n\to\infty} \frac{y_n - y_0}{n} = 0 \tag{42}$$

with probability 1. Since $y_t - y_s$ has a symmetric distribution with median 0 for every s, t, if $Q_n(l)$ is the function of concentration of $y_{n+1} - y_n$, the generalization of Lemma 4 to continuous processes gives

$$P\{\underset{0\leq t\leq 1}{\text{L.U.B.}} |y_{n+t} - y_n| > 2\epsilon n\} \leqq 2[1 - Q_n(\epsilon n)] \leqq 2P\{|y_{n+1} - y_n| > \tfrac{1}{2}\epsilon n\}. \tag{43}$$

Now because of (42), $|y_{n+1} - y_n| < \frac{1}{2}\epsilon n$ for sufficiently large n, with probability 1. Then, by a lemma of Borel, the terms of (43) are those of a convergent series,[27] so that

$$\sum_{n=0}^{\infty} P\{\underset{0\leq t\leq 1}{\text{L.U.B.}} |y_{n+t} - y_t| > 2\epsilon n\} < \infty, \tag{44}$$

i.e., by the converse of Borel's lemma,[28]

$$\underset{0\leq t\leq 1}{\text{L.U.B.}} \frac{|y_{n+t} - y_n|}{n} \leqq 2\epsilon, \tag{45}$$

for sufficiently large n, with probability 1. Thus

$$\limsup_{n\to\infty} \frac{1}{n} \underset{0\leq t\leq 1}{\text{L.U.B.}} |y_{n+t} - y_n| \leqq 2\epsilon, \tag{46}$$

with probability 1. Since ϵ can be made arbitrarily small,

$$\lim_{n\to\infty} \frac{1}{n} \underset{0\leq t\leq 1}{\text{L.U.B.}} |y_{n+t} - y_n| = 0 \tag{47}$$

with probability 1. Equations (42) and (47) imply together that

$$\lim_{t\to\infty} \frac{y_t - y_0}{t} = \lim_{t\to\infty} \left[\frac{x_t - x_0}{t} - \frac{\xi_t - \xi_0}{t}\right] = 0 \tag{48}$$

with probability 1. Then by Lemma 6, if E consists of the rational values of t, plus the set S_0, there is a numerically-valued function $f(t)$, defined on E, such that

$$\lim_{t\to\infty} \left[\frac{x_t - x_0}{t} - f(t)\right] = 0 \qquad (t \in E) \tag{49}$$

[27] Lévy, ibid., p. 126.

[28] Lévy, ibid., p. 127.

with probability 1. As in the proof of Lemma 6, we can suppose that $f(t)$ $(t \epsilon E)$ is a median value of $t^{-1}(x_t - x_0)$. It follows from the continuity properties of a centered differential process that $f(t)$ is bounded in every finite interval, so that if $t_0 \epsilon' E$, we can define $f(t_0)$ as $\overline{\lim\limits_{t \to t_0}} f(t)$ $(t \epsilon E)$, and $f(t)$ will then be a finite-valued function. We can assume that the space Ω of the process consists of functions $x(t)$ continuous except possibly for jumps (i.e., $\lim\limits_{t \uparrow t_0} x(t)$, $\lim\limits_{t \downarrow t_0} x(t)$ exist for all t_0).[29] If $x(t)$ is such a function,

$$\text{(50)} \qquad \underset{t>\tau}{\text{L.U.B.}} \left| \frac{x(t) - x(0)}{t} - f(t) \right| = \underset{\substack{t>\tau \\ t \epsilon E}}{\text{L.U.B.}} \left| \frac{x(t) - x(0)}{t} - f(t) \right|$$

so that (49) holds with no restriction on t, with probability 1.

THEOREM 9. *The preceding theorem remains true if the convergence in* (40) *and* (41) *is in probability, instead of with probability* 1.

The proof is essentially the same, except that (43) now can not be summed over n. By hypothesis, however, the right side of (43) approaches 0, as $n \to \infty$, so the left side does also, i.e.,

$$\text{(51)} \qquad \frac{1}{n} \underset{0 \leq t \leq 1}{\text{L.U.B.}} | y_{n+t} - y_n | \to 0$$

in probability. Then proceeding as before, we find that $t^{-1}(y_t - y_0) \to 0$ in probability so that, by Lemma 6, $f(t)$ can be chosen so that $t^{-1}(x_t - x_0) - f(t) \to 0$ in probability.

It is natural to ask when $E(x_t - x_0) = e(t)$ exists for all $t > 0$, if $f(t)$ can be taken as $te(t)$ in Theorems 8 and 9. The following theorem answers this question partially.

THEOREM 10. *Let* $\{x_t\}$ *be the variables of a centered differential process, and suppose that the chance variables* $\{x_n - x_{n-1} - E(x_n - x_{n-1})\}$, $n > 1$, *are uniformly integrable. Let* $E(x_t - x_0) = e(t)$. *Then*

$$\text{(52)} \qquad \frac{x_n - x_0 - e(n)}{n} \to 0$$

and

$$\text{(53)} \qquad \frac{x_t - x_0 - e(t)}{t} \to 0$$

in probability. If (52) *is true with probability* 1, (53) *is also.*

[29] Doob, Transactions of the American Mathematical Society, vol. 42(1937), pp. 134–135. Cf. also Lévy, Annali della Scuola Normale Superiore di Pisa, (2), vol. 3(1934), pp. 359–364, and vol. 4(1935), pp. 217–218.

It will be no restriction, because of Theorem 5, to assume that $e(t) \equiv 0$. The independent chance variables $\{x_n - x_{n-1}\}$ with 0 expectations are uniformly integrable, so

$$\frac{1}{N}\sum_{1}^{N}(x_n - x_{n-1}) = \frac{1}{N}(x_N - x_0) \to 0 \tag{54}$$

in probability, by a theorem of Persidskij.[30] Because of the uniform integrability, term by term integration is permissible: $N^{-1}E\,|\,x_N - x_0\,| \to 0$.[31] Now according to Lemma 2, $E\,|\,x_t - x_0\,| \leqq E\,|\,x_N - x_0\,|$ if $0 \leqq t \leqq N$, so $t^{-1}E\,|\,x_t - x_0\,| \to 0$, which implies that $t^{-1}(x_t - x_0) \to 0$ in probability. If (52) is true with probability 1, there is a function $e_1(t)$, according to Theorem 9, such that $t^{-1}[x_t - x_0 - e_1(t)] \to 0$ with probability 1. But since (53) is still true (with convergence in probability), $t^{-1}[e(t) - e_1(t)] \to 0$ so that $t^{-1}[x_t - x_0 - e(t)] \to 0$ with probability 1.

It is sufficient for uniform integrability in the above, if there is a positive number δ such that the expectations $E\{|\,x_n - x_{n-1} - E(x_n - x_{n-1})\,|^{1+\delta}\}$, $n \geqq 1$, exist and are uniformly bounded. In the temporally homogeneous case, δ can be taken $= 0$: integrability implies uniform integrability.

THEOREM 11. *Let $\{x_t\}$ be the variables of a centered temporally homogeneous differential process, and suppose that $E(x_1 - x_0) = e$ exists. Then*

$$\lim_{t\to\infty}\frac{x_t - x_0}{t} = e \tag{55}$$

with probability 1.

The hypotheses of Theorem 10 are satisfied, and moreover the convergence in (52) is with probability 1, since a set of independent chance variables $u_1, u_2, \cdots$ with identical distributions, whose expectations exist, obeys the strong law of large numbers.[32]

UNIVERSITY OF ILLINOIS.

[30] Comptes Rendus de l'Académie des Sciences de l'URSS, vol. 18(1938), p. 83.

[31] C. de la Vallée Poussin, Transactions of the American Mathematical Society, vol. 16(1915), pp. 445–447. We are using here the fact that uniform integrability of a sequence implies the uniform integrability of the corresponding sequence of averages.

[32] This fact is the analogue of Theorem 3 for a process depending on the parameter n instead of t, and has been proved by essentially the same method by Hopf (Journal of Mathematics and Physics of the Massachusetts Institute of Technology, vol. 13(1934), p. 92) and Doob (Transactions of the American Mathematical Society, vol. 36(1934), p. 764). Kolmogoroff (cf. M. Fréchet, *Recherches Théoriques Modernes sur la Théorie des Probabilités*, I: *Généralités sur les Probabilités. Variables Aléatoires*, pp. 257–259) has given an entirely different proof.

SUPERPOSITION ON MONOTONIC FUNCTIONS

By Esther McCormick Torrance

The object of this paper is the study of the superposition of a function on a general monotonic function $y = g(x)$ or on an inverse of such a function, $x = g^{-1}(y)$, and the associated problem of what properties the set A can have when A' has a stipulated property and $g(A') = A$ or $g^{-1}(A') = A$. The first part of the paper deals with the functions $h(x) = f[g(x)]$ and $k(x) = f[g^{-1}(x)]$ when restrictions are put on $f(y)$, and $g(x)$ is any monotonic function. The results obtained are contrasted with the results for superposition on continuous functions, and are used to determine relations between classes of sets. The second part of the paper deals with the function $h(x) = f[g(x)]$ when restrictions are put on $g(x)$. The results obtained include a generalization of a well-known theorem concerning such superposition and give a good example of the Baire-measure duality. The paper closes with an application of the results to the Stieltjes integral.

The functions studied in this paper have as their domain of definition and range of values the interval [0,1]. The terminology is that of Kuratowski [3].[1]

1. A monotonic function considered as a transformation consists of a homeomorphism between two G_δ's, H and H', the relation $g(K) = K'$, and the relation $g^{-1}(K') = K$, where K' is a denumerable set of points and K is the sum of a denumerable set of points and a denumerable set of closed intervals (see [11]). To every point of K' corresponds one point or one interval of K, but the interval can be open, semi-closed, or closed.

A property P of sets is said to be a *restricted intrinsic invariant property* if any set homeomorphic to a set having property P also has property P, and if whenever a set X has property P, $G_\delta X + F_\sigma$ also has property P.

Theorem 1. *A general monotonic transformation and its inverse carry sets having a restricted intrinsic invariant property* P *into sets having the same property* P.

Let X have a restricted intrinsic invariant property P. Then

$$g(X) = g(XH + XK) = g(XH) + g(XK) = X'H' + g(XK).$$

Since X has a restricted intrinsic invariant property P and H is a G_δ, XH has a restricted intrinsic invariant property P. Since the monotonic function $g(x)$ is a homeomorphism between H and H', it transforms XH into a set $X'H'$

Received June 13, 1939; presented in part to the American Mathematical Society, September 7, 1938.

[1] Numbers in brackets refer to the bibliography at the end of the paper.

homeomorphic to XH, and thus into a set with property P. Since $g(XK)$ must be a subset of K', it is at most a denumerable set of points, and so it is an F_σ. Hence $X'H' + g(XH)$ is the sum of a set having property P and an F_σ; thus it is a set having property P, so that $g(x)$ has property P.

Let X' have a restricted intrinsic invariant property P. Then

$$g^{-1}(X') = g^{-1}(X'H' + X'K') = g^{-1}(X'H') + g^{-1}(X'K').$$

The set $X'H'$ is the product of a set having the property P and a G_δ, and hence is a set with property P. Since $g^{-1}(X'H')$ is a homeomorph of $X'H'$, it must have property P. The set $g^{-1}(X'K')$ is a subset of K, so it is at most the sum of a denumerable set of points and a denumerable set of intervals, and hence is an F_σ. Then $g^{-1}(X')$ is the sum of a set having property P and an F_σ; hence it is a set having property P.

A linear set E is *perfectly measurable* if it is non-denumerable and if every set homeomorphic to E is measurable [4].

Let $E_1, E_2, \cdots$ be a denumerable sequence of sets on the x-axis. Let $r_1, r_2, \cdots$ be the rational points on the y-axis ordered in some arbitrary but fixed manner. In the XOY-plane erect perpendiculars to the y-axis through every point r_n and mark off on this line the points whose abscissas belong to E_n; call this set e_n. The *crible* C is the set $\sum e_n$. Let P_x be the line perpendicular to the x-axis at x, and R_x the set of points in $P_x(\sum e_n)$. Consider the set c of points ν such that R_ν has a subset of points $p_1, p_2, \cdots$ with $p_{k+1} \leqq p_k$ for every k. Call this set the positively cribled set by means of crible C. Call its complement the negatively cribled set.

An operation which permits us to pass from sets $E_1, E_2, \cdots$ to the positively and negatively cribled sets will be called the *crible operation*.

The crible sets C are the sets formed by the crible operation operating on intervals [5].

A family F_α or G_α of Borel sets is called a *multiplicative family of class* α if the product of a denumerable number of members of this family is still a member of this family.

A function $y = f(x)$ is a *Baire function of class* α if every closed set F in the space $\mathfrak{Y}$ of the range of values of the function is transformed by $f^{-1}(y) = x$ into a multiplicative Borel set of class α; i.e., for every closed set $F, f^{-1}(F) = A_F$, where A_F is a multiplicative Borel set of class α [3].

THEOREM 2. *The following classes of sets have a restricted intrinsic invariant property* P *as their defining property*: (1) *additive Borel sets of class* α, $\alpha > 1$; (2) *multiplicative Borel sets of class* α, $\alpha > 1$; (3) *sets having the Baire property in the restricted sense*; (4) *perfectly measurable sets*; (5) *projective sets of class n*.

The property distinguishing these classes is a property invariant under homeomorphisms (see [3], pp. 217 and 243, and [4]). Furthermore, sets G_δ and F_σ have each of these properties (1)–(5), and the class of sets having each of these properties is closed with respect to finite addition and multiplication,

so if X has any one of these properties, $G_\delta X + F_\sigma$ has it also. Hence each of these classes of sets has a restricted intrinsic invariant property P as its defining property.

In like manner it can be shown that crible sets C (see [5]) have a restricted intrinsic invariant property.

From Theorems 1 and 2 we can deduce the fact that all classes of sets listed in Theorem 2 are such that every monotonic transformation transforms a set of this class into a set of the same class. From this result we can sharpen for monotonic functions the well-known theorem which says that if $f(y)$ is a Baire function of class α and $g(x)$ is a Baire function of class 1, then $h(x) = f[g(x)]$ is a Baire function of class $\alpha + 1$.

THEOREM 3. *If $f(y)$ is a Baire function of class α, $\alpha > 1$, then $f[g(x)]$ is a Baire function of class α.*

Let $h(x) = f[g(x)]$. Then $h^{-1}(F) = g^{-1}[f^{-1}(F)] = g^{-1}(X)$, where X is a multiplicative Borel set of class α, $\alpha > 1$, and F is an arbitrary closed set. Since multiplicative Borel sets of class α, $\alpha > 1$, are transformed by the inverse monotonic function into sets of the same class, $g^{-1}(X)$ is a set of this class, so $h(x)$ must be a Baire function of class α.

A non-denumerable set of points E is said to be *concentrated in the neighborhood of a denumerable set H* if any open set containing the set H contains also the set E with the exception of at most a denumerable set of points [1].

If we confine ourselves to spaces in the interval [0, 1], we can give an equivalent definition of concentrated sets: a non-denumerable set C is a concentrated set if there exists a denumerable set $\{A(j)\}$ of increasing or decreasing sequences of points $A(j) = \{a_n(j)\}$ such that if for every j an open set G contains an interval I_j which contains an infinite number of points of the sequence $A(j)$, then the open set G contains all but a denumerable set of points of C [12]. If we use this definition, it is easy to show that the property of being a concentrated set is hereditary and invariant under monotonic homeomorphisms, and we have the following

LEMMA 1. *A monotonic transformation from space $\mathfrak{X}$ to space $\mathfrak{Y}$ transforms concentrated sets into concentrated sets or denumerable sets. A monotonic transformation from space $\mathfrak{Y}$ to space $\mathfrak{X}$ transforms concentrated sets into concentrated sets, or denumerable sets, plus open sets.*

LEMMA 2. *Every set having property* L *is a concentrated set.*

Let E be a set which has property L. Let $\sum x_n$ be a denumerable subset of E which is dense in E. Let G be any open set containing $\sum x_n$. Consider the set of points $A = (1 - G)E$. Set A is non-dense in E because in every set G_1 open in E with $G_1 E \neq 0$ there exists a point of the set $\sum x_n$ which is contained in the set G, so GG_1 is a set open in E and lying in the set G_1 and containing no point of A. Then set A is non-dense in the entire space, and so also is the closure of A. The closure of A is either a denumerable set of points B or the sum

of a perfect non-dense set P and a denumerable set of points B, $\bar{A} = P + B$. From the definition of A, A is the product of a closed set $(1 - G)$ and the set E. Then $A = \bar{A}E$. If $\bar{A}$ is the sum of a perfect non-dense set P and a denumerable set of points B, then $A = (P + B)E$. Since E has property L, it can have at most a denumerable number of points in common with the perfect non-dense set P, so that A is at most a denumerable set of points. Then E is a concentrated set.

A totally imperfect set is a set containing no perfect subset.

Lemma 3. *The property of being a totally imperfect set is an invariant of the monotonic transformation from the space $\mathfrak{X}$ to the space $\mathfrak{Y}$, and the monotonic transformation from the space $\mathfrak{Y}$ to the space $\mathfrak{X}$ carries a totally imperfect set into the sum of a totally imperfect set and an open set.*

Consider a totally imperfect set X. It is transformed into $H'X'$ by the homeomorphism associated with the monotonic function. If $H'X'$ contains any perfect subset, HX must contain a non-denumerable Borel set and hence a perfect subset. By definition, X does not contain any perfect subset. The set $(1 - H')X'$ consists of some of the points of K', $\sum p'_n$, so $X' = H'X' + \sum p'_n$ is totally imperfect since it contains no perfect subset.

In like manner consider a totally imperfect set X'. It is transformed into $HX + AK$, where HX is totally imperfect, and AK is at most the sum of a denumerable set of points and an open set.

From Lemma 2 we deduce the fact that monotonic functions transform sets having property L into concentrated sets or denumerable sets. It is easy to give an example of a monotonic function which does not transform a set having property L into a set having property L.

Sierpinski [7] has defined and proved the existence of a class of sets of power $\aleph_1$ whose homeomorphs are of measure zero and always of the first category. Let us call the class of all non-denumerable subsets of sets having this property the class of sets having property S. It can be shown very easily that a monotonic function transforms sets having property S into sets having this property, and an inverse monotonic function transforms any set having property S into the sum of a set of this type and an open set.

A set always of the first category is a set which is totally imperfect and which has the restricted Baire property ([3], p. 269).

Lemma 4. *The monotonic transformation from the space $\mathfrak{X}$ to the space $\mathfrak{Y}$ takes sets always of the first category into sets always of the first category, and the monotonic transformation from the space $\mathfrak{Y}$ to the space $\mathfrak{X}$ takes any set always of the first category into the sum of a set always of the first category and an open set.*

A totally imperfect set is transformed into a totally imperfect set by the monotonic transformation from the space $\mathfrak{X}$ to the space $\mathfrak{Y}$, and a totally imperfect set is transformed into the sum of a totally imperfect set and an open set by the monotonic transformation from the space $\mathfrak{Y}$ to the space $\mathfrak{X}$ by Lemma 3. A set having the restricted Baire property is transformed into a set with

the restricted Baire property by Theorems 1 and 2. It follows that a set having both these properties, that is, a set always of the first category, is transformed into a set always of the first category by the monotonic transformation from the space $\mathfrak{X}$ to the space $\mathfrak{Y}$, and a set always of the first category in the space $\mathfrak{Y}$ is transformed into the sum of a set always of the first category and an open set.

Collecting together these results, and noting that Theorems 1 and 2 imply that all the classes of sets mentioned in Theorem 2 are invariant under monotonic functions and their inverses, we obtain

THEOREM 4. *The following classes of sets have as their defining property a property invariant under every monotonic transformation from the space $\mathfrak{X}$ to the space $\mathfrak{Y}$ and from the space $\mathfrak{Y}$ to the space $\mathfrak{X}$*: (1) *projective sets of class n*; (2) *multiplicative Borel sets of class α, $\alpha > 1$*; (3) *additive Borel sets of class α, $\alpha > 1$*; (4) *sets having the restricted Baire property*; (5) *crible sets C*; (6) *perfectly measurable sets.*

The following classes of sets have as their defining property a property invariant under every monotonic transformation from the space $\mathfrak{X}$ to the space $\mathfrak{Y}$, and a monotonic transformation from the space $\mathfrak{Y}$ to the space $\mathfrak{X}$ transforms a set of one of these classes into the sum of a set of the same class and an open set: (1) *concentrated sets and denumerable sets,* (2) *totally imperfect sets,* (3) *sets with property* S *and denumerable sets,* (4) *sets always of the first category.*

For a large number of classes of sets we can find examples of members A and B of these classes and monotonic functions $g_1(x)$ and $g_2(x)$ for which $g_1(A) = A'$ does not give a member of the same class and $g_2^{-1}(B) = B'$ does not give a member of the same class or the sum of a member of such a class and an open set [12]. The classes for which such examples exist are listed in Theorem 5.

THEOREM 5. *The following classes of sets are classes which are not invariant under every monotonic transformation, whether the monotonic transformation is considered as a transformation from the space $\mathfrak{X}$ to the space $\mathfrak{Y}$ or as a transformation from the space $\mathfrak{Y}$ to the space $\mathfrak{X}$*: (1) *open sets,* (2) *closed sets,* (3) G_δ, (4) F_σ, (5) *dense sets,* (6) *frontier sets,* (7) *non-dense sets,* (8) *sets dense in themselves,* (9) *sets of the first category,* (10) *Baire sets,* (11) *sets with property* L *and denumerable sets.*

Let us compare these results with the results for continuous functions.

There exists a continuous function $c(x)$ and a projective set A of class n, where n is even, such that $c(A)$ is a projective set of class $n + 1$ ([3], p. 239). By Theorem 4 every monotonic function $g(x)$ transforms A into a projective set of class n.

Sierpinski [8] has shown that there exists a totally imperfect set A and there exists a continuous function $c(x)$ such that $c(A)$ is not totally imperfect. By Theorem 4, every monotonic function is such that $g(A)$ is totally imperfect when A is totally imperfect.

Sierpinski [9] has shown that there exists a continuous function $c(x)$ and a func-

tion $f(y)$ having the restricted Baire property which is such that $h(x) = f[c(x)]$ does not have the restricted Baire property. Suppose $f(y)$ has the restricted Baire property and consider $h(x) = f[g(x)]$, where $g(x)$ is a monotonic function. For every closed set F, $h^{-1}(F) = g^{-1}[f^{-1}(F)] = g^{-1}(A)$, where A has the restricted Baire property; so by Theorem 4, $g^{-1}(A)$ has the restricted Baire property and the function $h(x)$ has the restricted Baire property.

2. Let us consider the function $h(x) = f[g(x)]$ when restrictions are put on $g(x)$.

Lemma 5. *If a homeomorphism between H and H' is not absolutely continuous from H to H', it takes a point set of measure zero in H into a point set of positive measure in H'.*

Suppose a homeomorphism is not absolutely continuous from H to H'. Then given an arbitrary small positive quantity ϵ there is no δ such that $\mu(X) < \delta$ implies $\mu[f(X)] < \epsilon$; that is, for every δ there exists a set X such that $\mu(X) < \delta$ and $\mu[f(X)] > \epsilon$. Call the class of all the sets X such that $\mu(X) < \delta$ and $\mu[f(X)] > \epsilon$ the class X_δ. Every set in this class has measure less than δ. Then for any ϵ consider the ensemble $\{X_\delta\}$ of these classes X_δ. Consider some sequence $\{\delta_n\}$ of δ's which is such that $\sum \delta_n$ is finite. Take one set X_n from each of the classes X_{δ_n}. The measure of X_n is less than δ_n. Then the set $G_m = \sum_m^\infty X_n$ has a measure which is less than $\sum_m^\infty \delta_n$, and the measure of the set $\lim_{m\to\infty} G_m$ is zero since $\lim_{m\to\infty} \sum_m^\infty \delta_n$ is zero and the G_m form a decreasing sequence $G_1 \supset G_2 \supset \cdots$. Corresponding to the set $G_m = \sum_m^\infty X_n$ is a set G'_m which contains X'_m. Recall that X_m was a set with $\mu(X_m) < \delta$ and $\mu[f(X_m)] > \epsilon$, so X'_m has a measure greater than ϵ. Then $G'_1 \supset G'_2 \supset \cdots$ is a decreasing sequence of sets each having measure greater than ϵ, so $\lim_{m\to\infty} G'_m$ has measure greater than or equal to ϵ; hence it is a set of positive measure. Then a non-absolutely continuous transformation from H to H' takes a set of measure zero into a set of positive measure in H'.

Theorem 6. *A necessary and sufficient condition that $h(x) = f[g(x)]$ be measurable for every measurable function $f(y)$ is that the inverse monotonic function $x = g^{-1}(y)$ be absolutely continuous on H', the G'_δ associated with the monotonic function $y = g(x)$.*

The function $y = g(x)$ sets up a homeomorphism between H' in the space $\mathfrak{Y}$ and H in the space $\mathfrak{X}$. If $g^{-1}(y)$ is absolutely continuous as far as the homeomorphism set up by $y = g(x)$ is concerned on H', then given an arbitrary small positive quantity ϵ, there exists a δ such that if $\mu(Y) < \delta$ and $Y \subset H'$, then $\mu[g^{-1}(Y)] < \epsilon$, where $g^{-1}(Y)$ is a set in H since Y is in H'. Suppose we have a point set Z of measure zero in H'. Consider its transform $f(Z)$ in H. Since $\mu(Z) < \delta$, where δ is any positive number, then for every ϵ, $\mu[f(Z)] < \epsilon$, so the set $f(Z)$ is a point set of measure zero. Any measurable set A in the space $\mathfrak{Y}$

is composed of a Borel set B and a set of measure zero Z. Then $A = B + Z$ and $AH' = (B + Z)H' = BH' + ZH'$. The product of two Borel sets being a Borel set, BH' is a Borel set. The property of being a set of measure zero being hereditary, ZH' is of measure zero. Since the homeomorphism between H' and H takes every Borel set into a Borel set, BH' has as correspondent a Borel set B_1 in H. Since a set which is a Borel set with respect to a Borel set in a metric separable compact space is a Borel set in this space, B_1 is a Borel set in the space in question. The homeomorphism between H' and H carries point sets of measure zero in H' into point sets of measure zero, so ZH' is carried into a point set of measure zero, Z'. Then the part of A lying in H' is transformed into the sum of a Borel set and a set of measure zero in the space $\mathfrak{X}$ which is a measurable set in the space $\mathfrak{X}$.

Any point of A not lying in H' will either have no corresponding point in the space $\mathfrak{X}$ or will correspond to a point in the space $\mathfrak{X}$ or will correspond to an interval in the space $\mathfrak{X}$. There will be at most a denumerable number of such points and intervals and since the sum of a denumerable number of points and intervals is measurable, the part of A lying outside of H' is transformed into a measurable set in the space $\mathfrak{X}$.

Since $A = AH' + A(1 - H')$ is the sum of two sets both of which are transformed into measurable sets in the space $\mathfrak{X}$, A is transformed into a measurable set in the space $\mathfrak{X}$. Then if a set is measurable, it is transformed into a measurable set by any inverse monotonic function $x = g^{-1}(y)$ which is absolutely continuous on the G'_δ associated with the monotonic function $y = g(x)$. Consider any measurable function $f(y)$. The condition that $h(x) = f[g(x)]$ be measurable is that $h^{-1}(F)$ be measurable for every closed set F. Then $h^{-1}(F) = g^{-1}[f^{-1}(F)] = g^{-1}(A)$, where A is a measurable set. For functions $y = g(x)$ satisfying our conditions we have just proved that $g^{-1}(A)$ is measurable, so our conditions are sufficient that $f[g(x)]$ be measurable.

Now to prove that the conditions are necessary. If the homeomorphism between H' and H associated with $y = g(x)$ is not absolutely continuous from H' to H, then it takes a point set of measure zero, X', in H' into a point set of positive measure, X, in H, by Lemma 5. Consider some non-measurable subset of X, call it N. This corresponds to some subset $N'X'$ of X' which must be of measure zero since X' is. Then the transformation $y = g(x)$ takes a measurable set (here $N'X'$) into a non-measurable set. Consider the characteristic function $f(y)$ of $N'X'$. This is a measurable function since $N'X'$ is measurable. Take F as the closed set from $\frac{1}{2}$ to 1. Then

$$h^{-1}(F) = g^{-1}[f^{-1}(F)] = g^{-1}(N'X') = g^{-1}(H'N'X') + g^{-1}(N'X'K') = N + DK.$$

We thus have the sum of two disjoint sets, one a non-measurable set and the other a measurable set, so that the sum is non-measurable. Hence if the homeomorphism between H and H' associated with $y = g(x)$ is not absolutely continuous from H' to H, the compounded function $f[g(x)]$ is not measurable for every measurable function $f(y)$. This completes the proof.

This theorem is a generalization of the well-known theorem: in order that a

strictly increasing continuous function $y = g(x)$ be such that $h(x) = f[g(x)]$ is measurable for every measurable function $f(y)$ it is necessary and sufficient that $g^{-1}(y)$ be absolutely continuous.

THEOREM 7. *A necessary and sufficient condition that $h(x) = f[g(x)]$ be a measurable function for every function $f(y)$, where $g(x)$ is a monotonic function, is that the G_δ over which the function $y = g(x)$ defines a homeomorphism be of measure zero.*

First, we prove that if the G_δ set H associated with $g(x)$ is of measure zero, every set is carried into a measurable set by the monotonic transformation from the space $\mathfrak{Y}$ to the space $\mathfrak{X}$ associated with $g(x)$, and any function $f(y)$ is such that $h(x) = f[g(x)]$ is measurable. Let X' be any set and let X be its correspondent under the monotonic transformation. Then $g^{-1}(X') = g^{-1}(H'X' + X'K') = g^{-1}(H'X') + g^{-1}(X'K') = HX + AK = X$. Then X is a measurable set, being the sum of a set HX of measure zero and a Borel set.

Consider any function $f(y)$. Then $h^{-1}(F) = g^{-1}[f^{-1}(F)] = g^{-1}(X')$, X' being any set. We have just shown that $g^{-1}(X') = X$ is a measurable set, so $h(x) = f[g(x)]$ is measurable.

Now we prove that if every function $f(y)$ is such that $h(x) = f[g(x)]$ is measurable, where $g(x)$ is a monotonic function, then the G_δ set H over which $y = g(x)$ defines a homeomorphism is of measure zero. Suppose the G_δ set H is not of measure zero. Then it contains a non-measurable set M. Let M' be the transform of M by the monotonic homeomorphism defined by $y = g(x)$. Let $f(y)$ be the characteristic function of M', taking on the value $\frac{1}{2}$ at the points $y \in M'$, and zero elsewhere. Then $h^{-1}(\frac{1}{2}) = g^{-1}[f^{-1}(\frac{1}{2})] = g^{-1}(M') = M$ is not a measurable set, so $h(x)$ is not a measurable function. Then if H is not of measure zero, there exist functions $h(x) = f[g(x)]$ which are not measurable, and the hypothesis is contradicted. This completes the proof.

THEOREM 8. *A necessary and sufficient condition that the function $h(x) = f[g(x)]$ be non-measurable for every non-measurable function $f(y)$ is that the homeomorphism between the sets H and H' be absolutely continuous on H.*

Suppose the homeomorphism between H and H' is absolutely continuous on H. Then if $h(x)$ is a measurable function, for every closed set F, $h^{-1}(F) = g^{-1}[f^{-1}(F)] = g^{-1}(X') = X$ is a measurable set, so $X = B + Z$, where B is a Borel set and Z is a set of measure zero. Since the homeomorphism is absolutely continuous on H, $g(Z)$ is a set of measure zero. By Theorem 4, $g(B)$ is a Borel set. Then $f^{-1}(F) = X' = g(X)$ is a measurable set, so $f(y)$ must be a measurable function. Then if the homeomorphism between H and H' is absolutely continuous on H, $h(x)$ must be a non-measurable function if $f(y)$ is a non-measurable function.

The condition is also necessary, because if the homeomorphism between H and H' is not absolutely continuous, it transforms a set of measure zero into a set of positive measure by Lemma 5, and hence it transforms a set of measure zero into a non-measurable set. Suppose $f(y)$ is the characteristic function of this non-measurable set. Then $f(y)$ is non-measurable and $h(x)$ is measurable.

It is obvious that theorems similar to Theorems 6–8 could be proved for the function $k(x) = f[g^{-1}(x)]$.

THEOREM 9. *A necessary and sufficient condition that $f[g(x)] = h(x)$ have the Baire property whenever $f(y)$ has the Baire property is that the inverse function $g^{-1}(y)$ take sets in H' which are of the first category in $\mathfrak{Y}$ into sets of the first category in $\mathfrak{X}$.*

The condition is necessary. For, suppose it is not fulfilled. Then there is a set A' of the first category such that $g^{-1}(A'H') = A$ is not of the first category. Then there exists a subset of A, call it B, which does not have the Baire property. Now $g(B) = B'$ is a subset of A'. Let $f(y)$ be the characteristic function of B'. Then $f(y)$ has the Baire property since B' is a set of the first category. Now $h^{-1}(F) = g^{-1}[f^{-1}(F)] = g^{-1}(B') = B$ is a set not having the Baire property, so $h(x)$ does not have the Baire property.

The condition is sufficient. For $h^{-1}(F) = g^{-1}[f^{-1}(F)] = g^{-1}(A')$ where, since A' has the Baire property, $A' = B' + C'$, where B' is a Borel set and C' is a set of the first category. Since $g(x)$ is a Baire function of class one, $g^{-1}(B')$ is a Borel set B. Then $g^{-1}(B' + C') = g^{-1}(B') + g^{-1}(C') = B + g^{-1}[C'H' + C'K'] = B + C + DK$, where C is of the first category by hypothesis, and where DK is some denumerable set of points and denumerable set of intervals in K. Then $B + DK + C$ is the sum of a Borel set and a set of the first category, and is therefore a set with the Baire property.

THEOREM 10. *A necessary and sufficient condition that $h(x) = f[g(x)]$ have the Baire property for every function $f(y)$ is that H be a set of the first category.*

Suppose H is not a set of the first category. Then let X and $1 - X$ be a division of the interval into two totally imperfect sets; XH and $(1 - X)H$ are then totally imperfect sets. Suppose both of these sets are sets having the Baire property. Then both are the sum of a Borel set and a set of the first category. Since both are totally imperfect, the Borel set in both cases must be denumerable, and both XH and $(1 - X)H$ must be sets of the first category, so their sum H must be a set of the first category, and this is contrary to the hypothesis. Then there exists a subset of H which does not have the Baire property. Call this subset the set A, and call its transform A'. Let $f(y)$ be the characteristic function of A'. Then $f[g(x)] = h(x)$ is such that the closed set $F: [\frac{1}{2}, 1]$ has $h^{-1}(F) = g^{-1}[f^{-1}(F)] = g^{-1}(A') = g^{-1}(A'H) = A$, so $h(x)$ is not a Baire function. Then H must be a set of the first category in order that for every function $f(y)$ the function $f[g(x)]$ shall be a Baire function. The sufficiency is obvious.

THEOREM 11. *A necessary and sufficient condition that $f[g(x)] = h(x)$ fail to have the Baire property whenever $f(y)$ does not have the Baire property is that the function $g(x)$ take sets in H which are of the first category in the space $\mathfrak{X}$ into sets of the first category in the space $\mathfrak{Y}$.*

Suppose $f(y)$ does not have the Baire property. Then for some closed set F, $f^{-1}(F) = X'$ does not have the Baire property and $h^{-1}(F) = g^{-1}[f^{-1}(F)] = g^{-1}(X')$

$= g^{-1}(X'H') + g^{-1}(X'K') = XH + AK = X$, where AK is some denumerable set of points and intervals of K. If X has the Baire property and $g(x)$ takes sets in H of the first category in $\mathfrak{X}$ into sets of the first category in $\mathfrak{Y}$, then XH has the Baire property so that XH is the sum of a Borel set B and a set of the first category D, and $g(x) = g(B + D) = g(B) + g(DH) = A'K' = B' + D'$, and B' is a Borel set. Since $g(X) = X'$, $X' = B' + D'$ is a set having the Baire property. But X' is a set which does not have the Baire property, so the condition of the theorem is sufficient that X shall not have the Baire property, and hence that $h(x)$ shall be a function which fails to have the Baire property.

Suppose $g(x)$ takes a set X which is of the first category in $\mathfrak{X}$ and which is in H into a set which is not of the first category in $\mathfrak{Y}$. Then it takes some subset of this set into a set E which does not have the Baire property. Let $f(y)$ be the characteristic function of E. Then $f(y)$ does not have the Baire property and $h(x) = f[g(x)]$ does have the Baire property, so the condition of the theorem is necessary.

We note that Sierpinski considers measure and restricted Baire property as dual properties ([6], Chapter III), and studies applications of this duality. As Kondô and Szpilrajn [2, 10] point out, there is to some extent a duality between the restricted Baire property and perfect measure, and it is that duality and the duality between measure and the Baire property, between sets with perfect measure zero and sets always of the first category, that we have found useful, Theorems 9–11 being the duals of Theorems 6–8. It would be interesting to see how many of the theorems which Sierpinski did not dualize could be dualized by using this latter duality. For example, consider Sierpinski's theorem: *Every measurable function of a real variable transforms sets with property* $\mathfrak{S}$ *into sets always of the first category* ([6], p. 85).

A set has property $\mathfrak{S}$ if it has at most a denumerable set of points in common with every set of measure zero. The dual of a set with property $\mathfrak{S}$ is obviously a set with property L. Then the dual of Sierpinski's theorem is *Every function having the Baire property transforms sets with property* L *into sets of perfect measure zero.* This theorem I cannot prove, but it suggests a weaker theorem which is easy to prove.

THEOREM 12. *Every function having the Baire property transforms sets with property* L *into sets of measure zero.*

Let $\mathfrak{X}$, the domain of definition, be a complete metric separable space. If $f(x)$ has the Baire property, there exists a set Z such that $\mathfrak{X} - Z$ is of the first category and $f(Z)$ has measure zero. Let E be a set with property L. Then $f(E\mathfrak{X}) = f[E(\mathfrak{X} - Z) + EZ] = f[E(\mathfrak{X} - Z)] + f(EZ)$. The set $E(\mathfrak{X} - Z)$ is at most denumerable since E has property L, so $f(E\mathfrak{X}) = \sum p_i' + f(EZ)$ which must have measure zero since $f(Z)$ has measure zero.

3. Consider a Lebesgue-Stieltjes integral $\int f(x)\, dg(x)$, where $g(x)$ is a monotonic function. When does this integral exist?

Consider the substitution

$$y = g(x), \qquad x = g^{-1}(y), \qquad \int f(x)\, dg(x) = \int f[g^{-1}(y)]\, dy.$$

The integral exists when $k(y) = f[g^{-1}(y)]$ is measurable.

Szpilrajn has shown that a necessary and sufficient condition that $h(x) = f[g(x)]$ be measurable for every monotonic function $g(x)$ is that $f(x)$ be perfectly measurable, that is, $f^{-1}(F) = A$ be a perfectly measurable set for every closed set F. The same condition is necessary and sufficient for $k(y) = f[g^{-1}(y)]$ to be measurable. The sufficiency follows from Theorem 4. Associated with every monotonic function is a homeomorphism between H and H', and associated with this homeomorphism is an inverse monotonic function setting up the same homeomorphism, and vice versa. If a set is such that every inverse monotonic function transforms it into a measurable set, then every monotonic function must transform it into a measurable set, and hence the set must be perfectly measurable by Szpilrajn's theorem, so the condition is necessary. It is easy to show that if $f(x)$ is perfectly measurable, $k(y)$ is also, and vice versa.

We have noted that theorems similar to Theorems 6–8 hold for $k(y) = f[g^{-1}(y)]$. From these results we can say that

(1) A necessary and sufficient condition that the integral $\int f(x)\, dg(x)$ exists for every monotonic function $g(x)$ is that $f(x)$ be perfectly measurable (Szpilrajn).

(2) A necessary and sufficient condition that $\int f(x)\, dg(x)$ exist for every measurable function $f(x)$ and $g(x)$ a monotonic function is that $g(x)$ be absolutely continuous on H.

(3) A necessary and sufficient condition that $\int f(x)\, dg(x)$ exist for every function $f(x)$ and $g(x)$ a monotonic function is that H have measure zero.

Bibliography

1. A. S. Besicovitch, *Concentrated and rarefied sets of points*, Acta Math., vol. 62(1933), pp. 289–300.
2. M. Kondô, *Sur les notions de la catégorie*, Jour. Fac. Sc. Hokkaido Imp. Univ., Series I, Math., vol. IV, no. 3(1936), pp. 123–180.
3. K. Kuratowski, *Topologie* I, Monografje Matematyczne, Warsaw, 1933.
4. M. M. Lavrentieff, *Contribution à la théorie des ensembles homéomorphes*, Fund. Math., vol. 6(1924), pp. 149–160.
5. E. Selivanowski, *Sur une classe d'ensembles définis par une infinité dénombrable de conditions*, Comptes Rendus Paris, vol. 184(1927), pp. 1311–1313.
6. W. Sierpinski, *Hypothèse du Continu*, Warsaw, 1934.
7. W. Sierpinski, *Sur une extension de la notion de l'homéomorphie*, Fund. Math., vol. 22(1934), pp. 270–275.
8. W. Sierpinski, *Remarque sur le problème de l'invariance topologique de la propriété* (C), Fund. Math., vol. 30(1938), pp. 56–58.
9. W. Sierpinski, *Sur les fonctions jouissant de la propriété de Baire de fonctions continues*, Annals of Math., vol. 35(1934), pp. 278–283.
10. E. Szpilrajn, *Sur les ensembles et les fonctions absolument mesurables*, Comp. Rend. Soc. Sc. de Varsovie, vol. 30(1937), pp. 39–67.
11. E. M. Torrance, *Superposition on monotonic functions*, Fund. Math., vol. 30(1938), pp. 90–91.
12. E. M. Torrance, Thesis, Brown University, 1939.

Brown University.

THE CONVERGENCE OF EXPANSIONS RESULTING FROM A SELF-ADJOINT BOUNDARY PROBLEM

BY A. S. GALBRAITH AND S. E. WARSCHAWSKI

1. **Introduction.** This paper treats a problem of Riesz-Fischer type concerning the coefficients of the expansion of a function with n derivatives in terms of the characteristic solutions of a self-adjoint boundary problem of the second order. A corollary gives a criterion for term-by-term differentiation of such an expansion which is useful in certain applications.

The self-adjoint problem of this paper deals with the equation

$$(1) \qquad L(y) + \lambda r(x)y \equiv \frac{d}{dx}\left(p(x)\frac{dy}{dx}\right) - q(x)y + \lambda r(x)y = 0$$

containing the parameter λ, and the boundary conditions

$$(2) \qquad \begin{aligned} U_1(y, y') &\equiv \alpha_1 y(a) + \alpha_2 y(b) + \alpha_3 y'(a) + \alpha_4 y'(b) = 0, \\ U_2(y, y') &\equiv \beta_1 y(a) + \beta_2 y(b) + \beta_3 y'(a) + \beta_4 y'(b) = 0 \end{aligned}$$

subject to the following hypotheses:

(A) *$p(x)$, $p'(x)$, $q(x)$ and $r(x)$ are real continuous functions of the real variable x $(a \leqq x \leqq b)$, and $p(x)$ and $r(x)$ are positive in this closed interval.*

(B) *α_i, β_i $(i = 1, 2, 3, 4)$ are real constants, not all the determinants $\delta_{ij} \equiv \alpha_i\beta_j - \alpha_j\beta_i$ are zero, and*[1]

$$(3) \qquad p(a)\delta_{24} = p(b)\delta_{13}.$$

The theorems of the paper deal with a positive integer n, the number of times it is desired to differentiate the expansion of the function. It is assumed that

(C) *if $n > 2$, then $p(x)$, $q(x)$ and $r(x)$ have $n - 1$, $n - 2$ and $n - 2$ continuous derivatives, respectively* $(a \leqq x \leqq b)$.

It is known that for a denumerable set of values of λ solutions of (1) exist satisfying (2).[2] Such solutions $\varphi_k(x)$ $(k = 0, 1, 2, \cdots)$ are called *characteristic solutions* corresponding to the *characteristic numbers* $\lambda = \lambda_k$. *They are orthogonal and can be normalized*[3] so that

$$\int_a^b \varphi_i(x)\varphi_j(x)r(x)\,dx = \begin{cases} 0, & i \neq j; \\ 1, & i = j. \end{cases}$$

Received July 7, 1939; presented in part to the American Mathematical Society, April 7, 1939. The authors are indebted to Professor J. D. Tamarkin for suggesting the extension to the self-adjoint problem of their original results on Sturm-Liouville series.

[1] E. L. Ince, *Ordinary Differential Equations*, London, p. 216. This book will be referred to hereafter as Ince.

[2] See §2.

[3] Ince, pp. 237–238.

The principal results of the paper are the following.

THEOREM 1. *Let* $\{a_k\}$ *be a sequence of real constants such that* $\sum_{k=0}^{\infty} a_k^2 \lambda_k^n$ *converges. Then there exists a function* $f(x) \equiv \sum_{k=0}^{\infty} a_k \varphi_k(x)$ *such that* $\{a_k\}$ *are the Fourier coefficients of* $f(x)$, *i.e.,* $\int_a^b fr\varphi_k \, dx = a_k$; $f^{(n-1)}(x)$ *is absolutely continuous, and* $\int_a^b [f^{(n)}(x)]^2 \, dx$ *exists.*

COROLLARY. *The first* $n - 1$ *derivatives of* $f(x)$ *are given by*

$$f^{(s)}(x) = \sum_{k=0}^{\infty} a_k \varphi_k^{(s)}(x) \qquad (s = 0, 1, 2, \cdots, n-1),$$

the series converging absolutely and uniformly, and

$$\lim_{N\to\infty} \int_a^b [f^{(n)}(x) - \sum_{k=0}^{N} a_k \varphi_k^{(n)}(x)]^2 \, dx = 0.$$

THEOREM 2. *Let* $f(x)$ *be a function satisfying the following hypotheses* $(a \leqq x \leqq b)$:

(a) $f^{(n-1)}(x)$ *is absolutely continuous, and* $\int_a^b [f^{(n)}(x)]^2 \, dx$ *exists.*

(b) *If functions* $G_s(x)$ *are defined by the relations*

$$(1.1) \qquad \begin{aligned} G_0(x) &\equiv f(x), \\ -r(x)G_s(x) &\equiv L(G_{s-1}(x)) \qquad (s = 1, 2, \cdots, [\tfrac{1}{2}n] = m),\end{aligned}$$ [4]

the functions $G_0(x), G_1(x), \cdots, G_{m-1}(x)$ *satisfy the boundary conditions* (2), *if* $n > 1$. *If* $n = 1$, $f(x)$ *satisfies any boundary conditions* $y(a) = 0$, $y(b) = 0$, $y(a) = Ky(b)$ *implied by* (2), *and if* $n > 1$ *and* n *is odd,* $G_m(x)$ *satisfies any such boundary conditions.*[5]

Then if $a_k = \int_a^b fr\varphi_k \, dx$, *the series* $\sum_{k=0}^{\infty} a_k^2 \lambda_k^n$ *converges.*

Combination of the corollary to Theorem 1 with Theorem 2 gives

THEOREM 3. *If* $f(x)$ *satisfies the hypotheses of Theorem* 2 (which imply that $p(x)$, $q(x)$ and $r(x)$ have continuous derivatives of orders $n - 1$, $n - 2$ and $n - 2$ respectively if $n > 1$), *then*

$$f^{(s)}(x) = \sum_{k=0}^{\infty} a_k \varphi_k^{(s)}(x) \qquad (s = 0, 1, \cdots, n-1),$$

[4] $[\tfrac{1}{2}n]$ means the largest integer $\leqq \tfrac{1}{2}n$.

[5] These hypotheses are necessary conditions. They are satisfied for any function $f(x) = \sum_{k=0}^{\infty} a_k\varphi_k(x)$ for which $\sum_{k=0}^{\infty} a_k^2\lambda_k^n$ converges, as will be shown later. See footnote 15.

the convergence being absolute and uniform, and

$$\lim_{N\to\infty}\int_a^b [f^{(n)}(x) - \sum_{k=0}^{N} a_k\varphi_k^{(n)}(x)]^2\,dx = 0.$$

If $n = 1$, the hypotheses on the $G_s(x)$ are automatically satisfied, and *Theorem 3 reduces to an expansion theorem for $f(x)$ requiring only continuity of $q(x)$ and $r(x)$.*

For the general self-adjoint problem involving a differential equation of order $2m$ Krein[6] has given by different methods a theorem which for $m = 1$ becomes Theorems 1 and 2 of this paper with $n = 1$. While Krein's paper deals with a more general equation, the results of this paper for $n = 2, 3, \cdots$ are not included in his article. His proof uses the differentiability of the Green's function, and hence, if $m = 1$, does not seem directly applicable to the cases $n > 1$ treated in this paper.

2. Preliminaries.

(a) *Existence of characteristic numbers.* It was shown by Ettlinger[7] that under hypotheses (A) and (B) *the characteristic numbers* of the problem (1), (2) *form an increasing sequence* $\lambda_0 < \lambda_1 < \cdots < \lambda_n < \cdots, \lambda_n \to \infty$. Hence by writing (1) in the form

$$(py')' + [(\lambda - \lambda_0 + 1)r - (q - r(\lambda_0 - 1))]y = 0$$

the characteristic numbers $\Lambda_i = \lambda_i - \lambda_0 + 1$ are made $\geqq 1$. It will be assumed that (1) has been arranged in this way.

It is assumed that if two linearly independent characteristic solutions exist for any characteristic number,[8] they have been replaced by suitable linear combinations of themselves so as to be orthogonal. In this case the corresponding characteristic number will appear twice in the sequence, so that $\lambda_0 \leqq \lambda_1 \leqq \lambda_2 \leqq \cdots$.

(b) *Expansion of the Green's function.* If $G(x, t)$ is the Green's function[9] of the equation

$$(py')' - qy = 0$$

and the boundary conditions (2), then

$$\varphi_k(x) = -\lambda_k \int_a^b G(x, t)r(t)\varphi_k(t)\,dt, \tag{2.1}$$

[6] M. Krein, Recueil Math. de Moscou, new series, vol. 2(1937), p. 923.

[7] H. J. Ettlinger, *Existence theorems for the general real self-adjoint linear system of the second order*, Trans. Am. Math. Soc., vol. 19(1918), p. 94.

Cf. also E. Kamke, *Neue Herleitung der Oszillationssätze* $\cdots$, Math. Zeitschrift, vol. 44(1938–39), p. 654.

Under certain hypotheses of differentiability on the coefficients of the differential equation the existence of characteristic numbers and functions for a self-adjoint problem of order $2m$ was known earlier. See G. D. Birkhoff, Trans. Am. Math. Soc., vol. 9(1908), p. 373, and J. D. Tamarkin, Math. Zeitschrift, vol. 27(1928), p. 1.

[8] That not more than two can exist follows from the fact that the differential equation is of the second order.

[9] For the definition and properties of the Green's function see Ince, Chapter 11, p. 254.

or, if $(r(x))^{\frac{1}{2}}\varphi_k(x) = \mu_k(x)$ and $H(x, t) = G(x, t)(r(x))^{\frac{1}{2}}(r(t))^{\frac{1}{2}}$,

$$\mu_k(x) = -\lambda_k \int_a^b H(x, t)\mu_k(t)\, dt.$$

$H(x, t)$ is the continuous and symmetric kernel of this integral equation, of which all characteristic numbers are of the same sign, by subsection (a) above. It follows therefore from Mercer's theorem[10] that

$$(2.2) \qquad -H(x, t) = \sum_{k=0}^{\infty} \frac{\mu_k(x)\mu_k(t)}{\lambda_k} \qquad \text{or} \qquad -G(x, t) = \sum_{k=0}^{\infty} \frac{\varphi_k(x)\varphi_k(t)}{\lambda_k},$$

the convergence being absolute and uniform in both variables.

(c) *The set of characteristic functions* $\{\varphi_k(x)\}$ *is complete,*[11] *i.e., if* $f(x)$ *is a continuous function and* $a_k = \int_a^b fr\varphi_k\, dx$, *then*

$$\lim_{N\to\infty} \int_a^b [f(x) - \sum_{k=0}^{N} a_k\varphi_k(x)]^2\, dx = 0.$$

The proof uses

Lemma 1. *Given a function* $f(x)$ *continuous* $(a \leqq x \leqq b)$ *and given* $\epsilon > 0$ *there exists a function* $g(x)$ *continuous with its first two derivatives* $(a \leqq x \leqq b)$ *and satisfying the boundary conditions* (2) *such that* $\int_a^b [f(x) - g(x)]^2 dx < \epsilon$.

Proof. Let M be the least upper bound of $|f(x)|$ $(a \leqq x \leqq b)$. If δ is a positive number $< \frac{1}{2}(b - a)$ and $< \epsilon/(36M^2)$, there exists a polynomial $P(x)$ such that

$$\int_{a+\delta}^{b-\delta} [f(x) - P(x)]^2 dx < \tfrac{1}{2}\epsilon \quad \text{and} \quad |f(x) - P(x)| < \tfrac{1}{2}\epsilon \qquad (a + \delta \leqq x \leqq b - \delta).$$

Let c_a, c_a', c_b, c_b' be four constants each less than M in absolute value and such that

$$\alpha_1 c_a + \alpha_2 c_b + \alpha_3 c_a' + \alpha_4 c_b' = 0,$$

$$\beta_1 c_a + \beta_2 c_b + \beta_3 c_a' + \beta_4 c_b' = 0.$$

It is easily seen geometrically that there exist two functions $g_a(x)$ and $g_b(x)$ with the following properties: (i) $g_a(x)$ is continuous with its first two derivatives $(a \leqq x \leqq a + \delta)$; (ii) $g_a(a) = c_a$, $g_a'(a) = c_a'$; (iii) $g_a^{(m)}(a + \delta) = P^{(m)}(a + \delta)$

[10] J. Mercer, *Functions of positive and negative type and their connection with the theory of integral equations*, Phil. Trans. Royal Soc., (A), vol. 209(1909), p. 444. Kneser, *Integralgleichungen*, Braunschweig, p. 93.

[11] Under certain hypotheses of differentiability on the coefficients of (1) the completeness of the set $\{\varphi_k(x)\}$ was proved by Birkhoff and by Tamarkin, loc. cit. (see footnote 7). If $q(x)$ and $r(x)$ are merely continuous, the result can be inferred by the use of Lemma 1 or its equivalent from a theorem of Bliss, Trans. Am. Math. Soc., vol. 28(1926), p. 576. In the present paper a different proof is given which may be of interest.

($m = 0, 1, 2$); (iv) $g_b(x)$ has the corresponding properties in the interval $b - \delta \leqq x \leqq b$; (v) $| g_a(x) |, | g_b(x) | < 2M$. Let

$$g(x) = \begin{cases} g_a(x), & a \leqq x \leqq a + \delta, \\ P(x), & a + \delta \leqq x \leqq b - \delta, \\ g_b(x), & b - \delta \leqq x \leqq b. \end{cases}$$

Then $g(x)$ is continuous with its first two derivatives and satisfies the boundary conditions, and

$$\int_a^b [f(x) - g(x)]^2\,dx = \int_a^{a+\delta} [f(x) - g_a(x)]^2\,dx + \int_{a+\delta}^{b-\delta} [f(x) - P(x)]^2\,dx$$
$$+ \int_{b-\delta}^b [f(x) - g_b(x)]^2\,dx < 9M^2\delta + \tfrac{1}{2}\epsilon + 9M^2\delta.$$

Then since δ is chosen $< \epsilon/(36M^2)$, the result follows.

Proof of the completeness. By Lemma 1 a sequence of functions $\{g_m(x)\}$ exists, each continuous with its first two derivatives, each satisfying the boundary conditions (2) and such that

$$\lim_{m\to\infty} \int_a^b [f(x) - g_m(x)]^2\,dx = 0.$$

Let the functions $h_m(x)$ be defined by

$$(pg_m')' - qg_m = h_m r \qquad (m = 1, 2, \cdots).$$

Then[12]

$$\begin{aligned} g_m(x) &= \int_a^b G(x,t)h_m(t)r(t)\,dt \\ &= -\int_a^b \sum_{k=0}^{\infty} \frac{\varphi_k(x)\varphi_k(t)}{\lambda_k} h_m(t)r(t)\,dt \qquad \text{by (2.2)} \\ &= -\sum_{k=0}^{\infty} \frac{H_k^{(m)}\varphi_k(x)}{\lambda_k}, \end{aligned}$$

where $H_k^{(m)} = \int_a^b \varphi_k(t)h_m(t)r(t)\,dt$. The uniform convergence of the series for $G(x, t)$ justifies the integration term by term. Given $\epsilon > 0$, a number N exists such that

$$\int_a^b [f(x) - g_N(x)]^2\,dx < \tfrac{1}{4}\epsilon,$$

and for this N a number M exists such that

$$\left| g_N(x) + \sum_{k=0}^{M} \frac{H_k^{(N)}}{\lambda_k} \varphi_k(x) \right| < \left(\frac{\epsilon}{4(b - a)}\right)^{\frac{1}{2}} \qquad (a \leqq x \leqq b),$$

[12] Ince, p. 256.

because of the uniform convergence of the series. Then

$$\int_a^b \left[f(x) + \sum_{k=0}^{M} \frac{H_k^{(N)}}{\lambda_k} \varphi_k(x)\right]^2 dx \leqq 2\int_a^b [f(x) - g_N(x)]^2\, dx$$
$$+ 2\int_a^b \left[g_N(x) + \sum_{k=0}^{M} \frac{H_k^{(N)}}{\lambda_k} \varphi_k(x)\right]^2 dx < 2\cdot\frac{\epsilon}{4} + \frac{2\epsilon(b-a)}{4(b-a)} = \epsilon.$$

But if $a_k = \int_a^b f(t)r(t)\varphi_k(t)\, dt$, it is known that

$$\int_a^b \left[f(x) - \sum_{k=0}^{M} a_k\varphi_k(x)\right]^2 dx \leqq \int_a^b \left[f(x) + \sum_{k=0}^{M} \frac{H_k^{(N)}}{\lambda_k} \varphi_k(x)\right]^2 dx < \epsilon.$$

Hence the system of characteristic functions is complete.

Remark. An important consequence of the completeness of the set $\{\varphi_k(x)\}$ is that if the series $\sum_{k=0}^{\infty} \varphi_k(x) \int_a^b f(t)r(t)\varphi_k(t)\, dt$ converges uniformly, it converges to $f(x)$.

3. **Lemmas.** The following lemmas will be used in the proof of Theorem 1.

LEMMA 2. *Let $F(x)$ and $p(x)$ be continuous functions ($a \leqq x \leqq b$) and let $p(x)$ be positive in this closed interval. Divide the interval into M parts by means of the points $x_0 = a < x_1 < x_2 < \cdots < x_M = b$. Let $\Delta_i x = x_{i+1} - x_i$, $\Delta_i F = F(x_{i+1}) - F(x_i)$, $p_i = \min p(x)$ ($x_i \leqq x \leqq x_{i+1}$), and*

$$S_M = \sum_{i=0}^{M-1} \left(\frac{\Delta_i F}{\Delta_i x}\right)^2 p_i \Delta_i x.$$

Then

$$S_M \leqq \overline{\lim_{\max \Delta_i x \to 0}} S_M,$$

where the upper limit is taken for all methods of subdivision for which $\max \Delta_i x \to 0$.

Proof. Let $S_{M'}$ denote the sum obtained from S_M by bisecting each interval (x_i, x_{i+1}) by the point ξ_i. A typical term of S_M can be written as

$$\left(\frac{\Delta_i F}{\Delta_i x}\right)^2 p_i \Delta_i x = \left[\frac{F(x_{i+1}) - F(\xi_i)}{2(x_{i+1} - \xi_i)} + \frac{F(\xi_i) - F(x_i)}{2(\xi_i - x_i)}\right]^2 p_i \Delta_i x$$
$$\leqq 2\left[\left(\frac{F(x_{i+1}) - F(\xi_i)}{2(x_{i+1} - \xi_i)}\right)^2 + \left(\frac{F(\xi_i) - F(x_i)}{2(\xi_i - x_i)}\right)^2\right] p_i \Delta_i x.$$

Let

$$p_{i1} = \min p(x) \quad (\xi_i \leqq x \leqq x_{i+1}); \qquad p_{i2} = \min p(x) \quad (x_i \leqq x \leqq \xi_i).$$

Then

$$\left(\frac{\Delta_i F}{\Delta_i x}\right)^2 p_i \Delta_i x \leqq \left(\frac{F(x_{i+1}) - F(\xi_i)}{x_{i+1} - \xi_i}\right)^2 p_{i1}(x_{i+1} - \xi_i) + \left(\frac{F(\xi_i) - F(x_i)}{\xi_i - x_i}\right)^2 p_{i2}(\xi_i - x_i).$$

Hence

$$S_M \leqq S_{M'}.$$

Let now $S_{M'}, S_{M_1'}, S_{M_2'}, \cdots$ be a sequence of sums each of which is obtained from the preceding by the above method of subdivision. Then

$$S_M \leqq S_{M'} \leqq \overline{\lim_{k\to\infty}} S_{M_k'} \leqq \overline{\lim_{\max \Delta_i x\to 0}} S_M.$$

LEMMA 3. *If $\{\varphi_k(x)\}$ are the characteristic solutions of the problem* (1), (2), *then*

$$\sum_{k=0}^{\infty} \frac{\varphi_k^2(x)}{\lambda_k} \leqq K_0, \qquad \sum_{k=0}^{\infty} \frac{(\varphi_k'(x))^2}{\lambda_k^2} \leqq K_1 \qquad (a \leqq x \leqq b), \tag{3.1}$$

where K_0 and K_1 are constants. If in addition $p(x)$, $q(x)$ and $r(x)$ have continuous derivatives of orders $n-1$, $n-2$ and $n-2$ respectively $(a \leqq x \leqq b)$, then

$$\sum_{k=0}^{\infty} \frac{(\varphi_k^{(s)}(x))^2}{\lambda_k^{s+1}} \leqq K_s \qquad (a \leqq x \leqq b; s = 2, 3, \cdots, n),$$

where the K_s are constants.

Proof. By (2.2)

$$G(x, t) = -\sum_{k=0}^{\infty} \frac{\varphi_k(x)\varphi_k(t)}{\lambda_k},$$

the series being absolutely and uniformly convergent in x and t. If $x = t$, the first of the inequalities (3.1) follows. Differentiation of (2.1) gives

$$\varphi_k'(x) = -\lambda_k \int_a^b G_x(x, t) r(t) \varphi_k(t)\, dt.$$

Then the $\{-\varphi_k'(x)/\lambda_k\}$ are the Fourier coefficients of $G_x(x, t)$, which is continuous except for a jump of $1/p(x)$ at $t = x$ and is therefore bounded and integrable. Hence

$$\sum_{k=0}^{\infty} \frac{(\varphi_k'(x))^2}{\lambda_k^2} = \int_a^b r(t)[G_x(x, t)]^2\, dt \leqq K_1.$$

In the case of the second derivative $p'(x)$ is continuous, so from (1)

$$p\varphi_k'' + p'\varphi_k' + (\lambda_k r - q)\varphi_k = 0. \tag{3.2}$$

Then

$$\frac{p\varphi_k''}{\lambda_k^{\frac{3}{2}}} = -\frac{p'\varphi_k'}{\lambda_k^{\frac{3}{2}}} - \frac{r\varphi_k}{\lambda_k^{\frac{1}{2}}} + \frac{q\varphi_k}{\lambda_k^{\frac{3}{2}}},$$

$$\frac{(p\varphi_k'')^2}{\lambda_k^3} \leqq 3\left[\frac{(p'\varphi_k')^2}{\lambda_k^3} + \frac{(r\varphi_k)^2}{\lambda_k} + \frac{(q\varphi_k)^2}{\lambda_k^3}\right].$$

The results already obtained show that

$$\sum_{k=0}^{\infty} \frac{(\varphi_k''(x))^2}{\lambda_k^3} \leqq K_2 \qquad (a \leqq x \leqq b).$$

Differentiation of (3.2) $m - 2$ times and division by $\lambda_k^{\frac{1}{2}(m+1)}$ give the general result by an induction.

LEMMA 4. *If $\delta_{34} = 0$, then either $\delta_{14} \neq 0$ or the boundary conditions reduce to*

$$\text{(3.3)} \qquad \begin{aligned} h_1 y(a) &= h_2 y'(a), \qquad h_1^2 + h_2^2 > 0, \\ y(b) &= 0. \end{aligned}$$

Proof. If $\delta_{14} = 0$, then $\alpha_3\delta_{14} - \alpha_1\delta_{34} = \alpha_4\delta_{13} = 0$; $\beta_3\delta_{14} - \beta_1\delta_{34} = \beta_4\delta_{13} = 0$.

Case (a). If not both α_4 and β_4 are zero, suppose $\alpha_4 \neq 0$. Then $\delta_{13} = 0$. Therefore $\delta_{24} = 0$ by (3), and $\alpha_3\delta_{24} - \alpha_2\delta_{34} = \alpha_4\delta_{23} = 0$; $\alpha_2\delta_{14} - \alpha_1\delta_{24} = \alpha_4\delta_{12} = 0$. Hence $\delta_{12} = \delta_{23} = 0$. This means that all six $\delta_{ij} = 0$, contrary to hypothesis (B). The same result is obtained if $\beta_4 \neq 0$.

Case (b). $\alpha_4 = \beta_4 = 0$. Then $\delta_{24} = 0 = \delta_{13}$, by hypothesis (3). If $\alpha_3 = 0$, $\delta_{13} = \alpha_1\beta_3 = 0$. If $\beta_3 = 0$ also, then either $\delta_{12} = 0$ or $y(a) = y(b) = 0$. If $\delta_{12} = 0$, then all $\delta_{ij} = 0$, contrary to hypothesis (B). If $y(a) = y(b) = 0$, the boundary conditions reduce to (3.3) with $h_2 = 0$. If $\beta_3 \neq 0$, then $\alpha_1 = 0$ and $U_1 \equiv \alpha_2 y(b) = 0$. $\alpha_2 \neq 0$, for otherwise all the $\delta_{ij} = 0$. Hence $U_3 \equiv \alpha_2 U_2 - \beta_2 U_1 \equiv \alpha_2\beta_1 y(a) + \alpha_2\beta_3 y'(a) = 0$; and $U_1 = 0$ and $U_3 = 0$ form a set of boundary conditions of the kind specified above. If $\alpha_3 \neq 0$, then $\beta_3 U_1 - \alpha_3 U_2 \equiv y(b)\delta_{23} + y(a)\delta_{13} = 0$. Since $\delta_{13} = 0$, $y(b)\delta_{23} = 0$. But if $\delta_{23} = 0$, $\alpha_1\delta_{23} - \alpha_2\delta_{13} = -\alpha_3\delta_{12} = 0$, so $\delta_{12} = 0$, contrary to hypothesis (B). Hence $y(b) = 0$. This and $U_3 = 0$ are the required conditions (3.3).

The boundary conditions therefore take one of three forms:

(i) *If $\delta_{34} \neq 0$, (2) can be solved to give*

$$\text{(3.4)} \qquad \begin{aligned} y'(a)\delta_{34} &= -[y(a)\delta_{14} + y(b)\delta_{24}], \\ y'(b)\delta_{34} &= y(a)\delta_{13} + y(b)\delta_{23}. \end{aligned}$$

(ii) *If $\delta_{34} = 0$ but $\delta_{14} \neq 0$, (2) can be solved to give*

$$\text{(3.5)} \qquad \begin{aligned} y(a)\delta_{14} &= -y(b)\delta_{24}, \\ y'(b)\delta_{14} &= -[y(b)\delta_{12} + y'(a)\delta_{13}]. \end{aligned}$$

(iii) *If $\delta_{34} = \delta_{14} = 0$, the conditions become*

$$\text{(3.6)} \qquad \begin{aligned} h_1 y(a) &= h_2 y'(a), \qquad h_1^2 + h_2^2 > 0, \\ y(b) &= 0. \end{aligned}$$

4. **Proof of Theorem** 1.

Case 1: $n = 1$. *It is assumed that a sequence of constants $\{a_k\}$ ($k = 0, 1, 2, \cdots$) exists such that $\sum_{k=0}^{\infty} a_k^2\lambda_k$ converges. The conclusion is that there exists a function*

$f(x)$ which is absolutely continuous ($a \leqq x \leqq b$), $\int_a^b (f'(x))^2\, dx$ exists, and $a_k = \int_a^b fr\varphi_k\, dx$. It will be shown first that the series $\sum_{k=0}^{\infty} a_k\varphi_k(x)$ is absolutely and uniformly convergent ($a \leqq x \leqq b$), and therefore defines a continuous function $f(x) \equiv \sum_{k=0}^{\infty} a_k\varphi_k(x)$.

By Lemma 3

$$\sum_{k=0}^{\infty} \frac{\varphi_k^2(x)}{\lambda_k} \leqq K_0, \quad \text{a constant} \qquad (a \leqq x \leqq b).$$

Then

$$\Big(\sum_{k=N}^{N+M} |a_k|\,|\varphi_k(x)|\Big)^2 = \left(\sum_{k=N}^{N+M} |a_k|\,\lambda_k^{\frac{1}{2}} \frac{|\varphi_k(x)|}{\lambda_k^{\frac{1}{2}}}\right)^2$$

$$\leqq \Big(\sum_{k=N}^{N+M} a_k^2\lambda_k\Big)\left(\sum_{k=N}^{N+M} \frac{\varphi_k^2(x)}{\lambda_k}\right) \leqq K_0\Big(\sum_{k=N}^{N+M} a_k^2\lambda_k\Big).$$

From the convergence of $\sum_{k=0}^{\infty} a_k^2\lambda_k$ the result follows.

To prove the remaining conclusions consider

$$F(x) \equiv F_{R,N}(x) \equiv f(x) - \sum_{k=0}^{R-1} a_k\varphi_k(x) - \sum_{k=N+1}^{\infty} a_k\varphi_k(x) = \sum_{k=R}^{N} a_k\varphi_k(x),$$

where R and N are non-negative integers, $N > R$. If the interval (a, b) is subdivided as in Lemma 2, and $p(x)$ is the function $p(x)$ of (1), write

$$S_M = \sum_{i=0}^{M-1} \left(\frac{\Delta_i F}{\Delta_i x}\right)^2 p_i\,\Delta_i x$$

as in Lemma 2. By Lemma 2

$$S_M \leqq \overline{\lim_{\substack{m\to\infty \\ \max \Delta_i x \to 0}}}\; S_m, \tag{4.1}$$

where S_m is any similar sum for an arbitrary subdivision of (a, b). To find this upper limit consider

$$S_m = \sum_{i=0}^{m-1}\left(\frac{\sum_{k=R}^{N} a_k\Delta_i\varphi_k}{\Delta_i x}\right)^2 p_i\,\Delta_i x$$

$$= \sum_{i=0}^{m-1} p_i \sum_{k=R}^{N}\sum_{l=R}^{N} a_k a_l \frac{\Delta_i\varphi_k}{\Delta_i x}\cdot\frac{\Delta_i\varphi_l}{\Delta_i x}\cdot\Delta_i x$$

$$= \sum_{k=R}^{N}\sum_{l=R}^{N} a_k a_l \sum_{i=0}^{m-1} p_i \frac{\Delta_i\varphi_k}{\Delta_i x}\cdot\frac{\Delta_i\varphi_l}{\Delta_i x}\cdot\Delta_i x.$$

Since $p(x)$ is continuous and $\varphi_k(x)$ has a continuous derivative,

$$\lim_{\substack{m\to\infty \\ \max \Delta_i x \to 0}} S_m = \sum_{k=R}^{N} \sum_{l=R}^{N} a_k a_l \int_a^b p\varphi_k'(x)\varphi_l'(x)\,dx. \tag{4.2}$$

Integration by parts and use of (1) give for the right member

$$\sum_{k=R}^{N} \sum_{l=R}^{N} a_k a_l \left\{ [p\varphi_k\varphi_l']_a^b - \int_a^b \varphi_k(q - \lambda_l r)\varphi_l\,dx \right\}. \tag{4.3}$$

The three forms of the boundary conditions listed after Lemma 4 can be considered separately. If (3.4) is true, the integrated terms in (4.3) become

$$\sum_{k=R}^{N} \sum_{l=R}^{N} a_k a_l[A\varphi_k(b)\varphi_l(b) + B\varphi_k(a)\varphi_l(a) + C\varphi_k(a)\varphi_l(b) + D\varphi_k(b)\varphi_l(a)],$$

where A, B, C and D are constants independent of k, l, R, N. If (3.5) is true they become

$$p(b)\varphi_k(b)\varphi_l'(b) - p(a)\varphi_k(a)\varphi_l'(a)$$

$$= -\frac{1}{\delta_{14}}[\varphi_k(b)\varphi_l'(a)(p(b)\delta_{13} - p(a)\delta_{24})] - \frac{\delta_{12}}{\delta_{14}}p(b)\varphi_k(b)\varphi_l(b).$$

The square bracket vanishes, by (3), so the integrated terms reduce to

$$\sum_{k=R}^{N} \sum_{l=R}^{N} a_k a_l A\varphi_k(b)\varphi_l(b).$$

If (3.6) holds, the integrated terms either vanish or take the form

$$\sum_{k=R}^{N} \sum_{l=R}^{N} a_k a_l B\varphi_k(a)\varphi_l(a).$$

Then if $\sum_{k=R}^{N} a_k\varphi_k(x)$ is set equal to $\psi_{R,N}(x)$, the equation (4.2) becomes in each case

$$\lim_{\substack{m\to\infty \\ \max \Delta_i x \to 0}} S_m = A\psi_{R,N}^2(b) + B\psi_{R,N}^2(a) + C\psi_{R,N}(a)\psi_{R,N}(b) + \sum_{k=R}^{N} a_k^2\lambda_k - \int_a^b q\psi_{R,N}^2(x)\,dx,$$

where A, B and C are constants, independent of R and N. By (4.1) S_M is not greater than this. If M and R are held fast, while $N \to \infty$,

$$\begin{aligned} \lim_{N\to\infty} S_M &= \sum_{i=0}^{M-1}\left(\frac{\Delta_i f}{\Delta_i x} - \sum_{k=0}^{R-1} a_k \frac{\Delta_i \varphi_k}{\Delta_i x}\right)^2 p_i\Delta_i x \\ &\leqq A\lim_{N\to\infty}\psi_{R,N}^2(b) + B\lim_{N\to\infty}\psi_{R,N}^2(a) + C\lim_{N\to\infty}\psi_{R,N}(a)\lim_{N\to\infty}\psi_{R,N}(b) \\ &\quad + \sum_{k=R}^{\infty} a_k^2\lambda_k - \lim_{N\to\infty}\int_a^b q\psi_{R,N}^2(x)\,dx. \end{aligned} \tag{4.4}$$

Since the series $\sum_{k=0}^{\infty} a_k\varphi_k(x)$ has been shown to be uniformly convergent $(a \leqq x \leqq b)$, $\lim_{N\to\infty} \psi_{R,N}(x) = \sum_{k=R}^{\infty} a_k\varphi_k(x) \equiv \Psi_R(x)$. Hence

$$\begin{aligned}\lim_{N\to\infty} S_M &\leqq A\Psi_R^2(b) + B\Psi_R^2(a) + C\Psi_R(a)\Psi_R(b) + \sum_{k=R}^{\infty} a_k^2\lambda_k - \int_a^b q\Psi_R^2\,dx \\ &= B_R,\end{aligned} \tag{4.5}$$

say. If R is chosen as *zero*, (4.4) and (4.5) show that

$$\sum_{i=0}^{M-1}\left(\frac{\Delta_i f}{\Delta_i x}\right)^2 p_i\Delta_i x \leqq B_0,$$

a constant, for all M. Then by a theorem of Hahn[13] $f(x)$ is absolutely continuous, $f'(x)$ exists almost everywhere and $\int_a^b (f'(x))^2\,dx$ exists. This completes the proof for $n = 1$.

To prove the corollary observe that if $R \neq 0$, (4.4) and (4.5) show that

$$\sum_{i=0}^{M-1}\left(\frac{\Delta_i f}{\Delta_i x} - \sum_{k=0}^{R-1} a_k\frac{\Delta_i\varphi_k}{\Delta_i x}\right)^2 p_i\Delta_i x \leqq B_R,$$

so that Hahn's theorem proves that

$$\int_a^b \left(f' - \sum_{k=0}^{R-1} a_k\varphi_k'\right)^2 dx \leqq \frac{B_R}{\min_{[a,\,b]} p(x)}.$$

$\lim_{R\to\infty} B_R = 0$, since $\lim_{R\to\infty} \Psi_R(x) = \lim_{R\to\infty}\sum_{k=R}^{\infty} a_k\varphi_k = 0$ uniformly $(a \leqq x \leqq b)$. Therefore the derived series converges in the mean to $f'(x)$.

Case 2: $n = 2$. By hypothesis $\sum_{k=0}^{\infty} a_k^2\lambda_k^2$ converges. The function

$$f(x) \equiv \sum_{k=0}^{\infty} a_k\varphi_k(x)$$

is continuous with its first derivative, for

$$\left(\sum_{k=N}^{N+M} |a_k|\,|\varphi_k'(x)|\right)^2 = \left(\sum_{k=N}^{N+M} |a_k|\,\lambda_k\cdot\frac{|\varphi_k'(x)|}{\lambda_k}\right)^2 \leqq \left(\sum_{k=N}^{N+M} a_k^2\lambda_k^2\right)\left(\sum_{k=N}^{N+M}\frac{(\varphi_k'(x))^2}{\lambda_k^2}\right).$$

By Lemma 3 the second parenthesis is $\leqq K_1$, a constant, and the convergence of $\sum_{k=0}^{\infty} a_k^2\lambda_k^2$ shows that the first parenthesis can be made arbitrarily small by choosing

[13] Hahn, Monatshefte für Math. u. Physik, vol. 23(1912), p. 172. Or cf. Hobson, *Theory of Functions of a Real Variable*, Cambridge, 1927, vol. 1, p. 672.

N sufficiently large. Hence the series $\sum_{k=0}^{\infty} a_k\varphi_k'(x)$ is absolutely and uniformly convergent, so that $f'(x)$ is continuous.

The function $F'(x) \equiv F'_{R,N}(x)$ is defined by

$$F'_{R,N}(x) \equiv f'(x) - \sum_{k=0}^{R-1} a_k\varphi_k'(x) - \sum_{k=N+1}^{\infty} a_k\varphi_k'(x) = \sum_{k=R}^{N} a_k\varphi_k'(x).$$

Then as before let

$$S'_M = \sum_{i=0}^{M-1} \left(\frac{\Delta_i F'}{\Delta_i x}\right)^2 p_i^2 \Delta_i x.$$

Then by an argument similar to that used in the previous case

$$S'_M \leqq \lim_{\substack{m\to\infty \\ \max \Delta_i x \to 0}} S'_m = \sum_{k=R}^{N} \sum_{l=R}^{N} a_k a_l \int_a^b p^2 \varphi_k'' \varphi_l'' \, dx.$$

Use of (1) and occasional interchange of the subscripts of summation k and l give

$$\lim_{\substack{m\to\infty \\ \max \Delta_i x \to 0}} S'_m = \int_a^b q^2 \left(\sum_{k=R}^{N} a_k\varphi_k\right)^2 dx - 2\int_a^b p'q\left(\sum_{k=R}^{N} a_k\varphi_k\right)\left(\sum_{l=R}^{N} a_l\varphi_l'\right) dx$$

$$(4.6) \qquad + \int_a^b (p')^2\left(\sum_{k=R}^{N} a_k\varphi_k'\right)^2 dx + \int_a^b r^2 \left(\sum_{k=R}^{N} a_k\lambda_k\varphi_k\right)^2 dx$$

$$- 2\int_a^b rq\left(\sum_{k=R}^{N} a_k\lambda_k\varphi_k\right)\left(\sum_{l=R}^{N} a_l\varphi_l\right) dx + 2\int_a^b rp'\left(\sum_{k=R}^{N} a_k\lambda_k\varphi_k\right)\left(\sum_{l=R}^{N} a_l\varphi_l'\right) dx.$$

The uniform convergence of $\sum_{k=0}^{\infty} a_k\varphi_k(x)$ and $\sum_{k=0}^{\infty} a_k\varphi_k'(x)$ shows that the first three terms of the right member of (4.6) are bounded. The fourth term can be written

$$\int_a^b r\left(\sum_{k=R}^{N} a_k\lambda_k r^{\frac{1}{2}}\varphi_k\right)^2 dx \leqq (\max r(x))\cdot\int_a^b \sum_{k=R}^{N} \sum_{l=R}^{N} a_k a_l \lambda_k\lambda_l r\varphi_k\varphi_l \, dx$$

$$= (\max r(x))\left(\sum_{k=R}^{N} a_k^2\lambda_k^2\right) < A, \text{ a constant,}$$

since $\sum_{k=0}^{\infty} a_k^2\lambda_k^2$ converges by hypothesis. By Schwarz' inequality the absolute value of the fifth term is

$$2\left|\int_a^b qr^{\frac{1}{2}}\left(\sum_{l=R}^{N} a_l\varphi_l\right)\left(\sum_{k=R}^{N} a_k\lambda_k r^{\frac{1}{2}}\varphi_k\right) dx\right|$$

$$\leqq 2\left[\int_a^b q^2 r\left(\sum_{l=R}^{N} a_l\varphi_l\right)^2 dx\cdot\int_a^b \sum_{k=R}^{N} \sum_{l=R}^{N} a_k a_l \lambda_k\lambda_l r\varphi_k\varphi_l \, dx\right]^{\frac{1}{2}},$$

and so is not greater than $B[\sum_{k=R}^{N} a_k^2\lambda_k^2]^{\frac{1}{2}}$. The sixth term can be treated similarly. Hence the theorem of Hahn already used shows (if $R = 0$) that $f'(x)$ is absolutely continuous, $f''(x)$ exists almost everywhere and $\int_a^b [f''(x)]^2\, dx$ exists. If $R \neq 0$, the repeated limit, as first N and then R becomes infinite, is zero, so that $\sum_{k=0}^{R-1} a_k\varphi_k''(x)$ converges in the mean to $f''(x)$ as $R \to \infty$.

Before we proceed with the proof for $n > 2$, it is convenient to prove

LEMMA 5. *If $f(x) \equiv \sum_{k=0}^{\infty} a_k\varphi_k(x)$, where $a_k = \int_a^b fr\varphi_k\, dx$, and if $\sum_{k=0}^{\infty} a_k^2\lambda_k^{t+1}$ converges, where t is an integer $\geqq 0$, and if hypothesis* (C) *is satisfied with $n = t$, then the functions $G_s(x)$ as defined by* (1.1) *exist and satisfy the relations*

$$G_s(x) = \sum_{k=0}^{\infty} a_k\lambda_k^s\varphi_k(x) \qquad (s = 0, 1, 2, \cdots, [\tfrac{1}{2}t]), \tag{4.7}$$

the series converging absolutely and uniformly $(a \leqq x \leqq b)$.

Proof. If $t = 0$, $G_0 = f(x)$.

If $t = m > 0$, assume the theorem true for $t = 0, 1, \cdots, m - 1$. If m is odd, $[\frac{1}{2}m] = [\frac{1}{2}(m-1)]$ and hence nothing is to be proved. If m is even, say $m = 2i$, then the $G_s(x)$ exist and satisfy (4.7) for $s = 0, 1, 2, \cdots, i - 1$, and it must be proved that if $\sum_{k=0}^{\infty} a_k^2\lambda_k^{2i+1}$ converges,

$$G_i(x) = \sum_{k=0}^{\infty} a_k\lambda_k^i\varphi_k(x),$$

the series converging absolutely and uniformly. Now

$$-rG_i(x) = L(G_{i-1}) = pG_{i-1}'' + p'G_{i-1}' - qG_{i-1}$$

if the right member exists. The formal series for $G_{i-1}'(x)$ and $G_{i-1}''(x)$ converge absolutely and uniformly, for

$$\left(\sum_{k=N}^{N+M} |a_k\lambda_k^{i-1}|\,|\varphi_k''|\right)^2 = \left(\sum_{k=N}^{N+M} |a_k|\,\lambda_k^{i+\frac{1}{2}}\cdot\frac{|\varphi_k''|}{\lambda_k^{\frac{3}{2}}}\right)^2$$
$$\leqq \left(\sum_{k=N}^{N+M} a_k^2\lambda_k^{2i+1}\right)\cdot\left(\sum_{k=N}^{N+M}\frac{(\varphi_k'')^2}{\lambda_k^3}\right).$$

By Lemma 3 the second parenthesis is uniformly bounded, and since $\sum_{k=0}^{\infty} a_k^2\lambda_k^{2i+1}$ converges, the first parenthesis can be made arbitrarily small by taking N sufficiently large. Then the formal series for $G_{i-1}''(x)$ converges absolutely and uniformly. A similar argument holds for the series for $G_{i-1}'(x)$. Since the series

for $G_{i-1}(x)$ converges absolutely and uniformly by hypothesis, the derived series converge to $G'_{i-1}(x)$ and $G''_{i-1}(x)$, respectively. Therefore

$$- rG_i = \sum_{k=0}^{\infty} a_k \lambda_k^{i-1} [p\varphi_k'' + p'\varphi_k' - q\varphi_k]$$

$$= - \sum_{k=0}^{\infty} a_k \lambda_k^{i-1} \cdot \lambda_k r \varphi_k$$

by (1), so

$$G_i = \sum_{k=0}^{\infty} a_k \lambda_k^i \varphi_k .$$

Case 3: $n > 2$.

(a) $n = 2m + 1$. The hypothesis is that

$$\sum_{k=0}^{\infty} a_k^2 \lambda_k^{2m+1} \tag{4.8}$$

converges. The function $f(x)$ is defined by $f(x) \equiv \sum_{k=0}^{\infty} a_k \varphi_k(x)$. Also

$$\begin{aligned} \Big(\sum_{k=N}^{N+M} | a_k | \, | \varphi_k^{(s)}(x) | \Big)^2 &= \left(\sum_{k=N}^{N+M} | a_k | \lambda_k^{\frac{1}{2}s+\frac{1}{2}} \cdot \frac{| \varphi_k^{(s)}(x) |}{\lambda_k^{\frac{1}{2}s+\frac{1}{2}}} \right)^2 \\ &\leqq \Big(\sum_{k=N}^{N+M} a_k^2 \lambda_k^{s+1}\Big) \left(\sum_{k=N}^{N+M} \frac{(\varphi_k^{(s)}(x))^2}{\lambda_k^{s+1}} \right). \end{aligned} \tag{4.9}$$

By Lemma 3 the second parenthesis is $\leqq K_s$, a constant, and (4.8) shows that the first parenthesis is arbitrarily small for N sufficiently large, provided that $s \leqq 2m$. Hence the series for $f(x)$ can be differentiated term by term $2m$ times, the resulting series being absolutely and uniformly convergent.

By Lemma 5 $G_s(x)$ exists and

$$G_s(x) = \sum_{k=0}^{\infty} a_k \lambda_k^s \varphi_k(x) \qquad (s = 0, 1, 2, \cdots, m),$$

the series converging absolutely and uniformly. Hence the Fourier coefficients of $G_m(x)$ are $a_k \lambda_k^m = B_k$. (4.8) shows that $\sum_{k=0}^{\infty} B_k^2 \lambda_k$ converges. Consequently Theorem 1 for $n = 1$ can be applied to $G_m(x)$ to show that $G_m(x)$ is absolutely continuous and $\int_a^b [G'_m(x)]^2 \, dx$ exists. But

$$\begin{aligned} - rG_1 &= (pG_0')' - qG_0 = (pf')' - qf, \\ - rG_1' &= r'G_1 + (pf')'' - (qf)' \\ &= (pf')'' - (qf)' - \frac{r'}{r} [(pf')' - qf], \\ - rG_2 &= (pG_1')' - qG_1 \\ &= - \left[\frac{p}{r} \{(pf')'' - (qf)'\} - \frac{pr'}{r^2} \{(pf')' - qf\} \right]' + \frac{q}{r} \{(pf')' - qf\}. \end{aligned}$$

Continuation of this process gives

$$(4.10)\qquad \begin{aligned} -rG_m &= (pG'_{m-1})' - qG_{m-1} \\ &= \frac{p^m}{r^{m-1}} f^{(2m)} + \mathfrak{L}(f, f', \cdots, f^{(2m-1)}), \end{aligned}$$

where $\mathfrak{L}(f, f', \cdots, f^{(2m-1)})$ is linear in the indicated derivatives of $f(x)$ with coefficients which are polynomials in $p(x)$, $q(x)$, $1/r(x)$ and their derivatives up to and including those of orders $2m - 1$, $2m - 2$ and $2m - 2$, respectively. $f^{(2m-1)}(x)$ is absolutely continuous since $f^{(2m)}(x)$ is continuous on account of the uniform convergence of its series, and $p(x)$ and $r(x)$ are nowhere zero, so (4.10) shows that $f^{(2m)}(x)$ also is absolutely continuous. This, with the hypotheses on $p(x)$, $q(x)$ and $r(x)$, shows that $\mathfrak{L}(f, \cdots, f^{(2m-1)})$ has a continuous derivative, and, since $G'_m(x)$ exists almost everywhere and its square is integrable, (4.10) shows that $\int_a^b [f^{(2m+1)}(x)]^2\, dx$ exists.

(b) $n = 2m$. The hypothesis is that $\sum_{k=0}^{\infty} a_k^2 \lambda_k^{2m}$ converges. Definition of $f(x) \equiv \sum_{k=0}^{\infty} a_k \varphi_k(x)$ as in (a) gives a series which can be differentiated term by term $2m - 1$ times. By Lemma 5 $G_s(x)$ exists and

$$G_s(x) = \sum_{k=0}^{\infty} a_k \lambda_k^s \varphi_k(x) \qquad (s = 0, 1, 2, \cdots, m - 1).$$

Also, the series for $G_{m-1}(x)$ can be differentiated term by term once. Hence $G'_{m-1}(x)$ is continuous. The coefficient

$$A_k \equiv \int_a^b G_{m-1}(x) r(x) \varphi_k(x)\, dx = a_k \lambda_k^{m-1},$$

so that

$$\sum_{k=0}^{\infty} A_k^2 \lambda_k^2 = \sum_{k=0}^{\infty} a_k^2 \lambda_k^{2m-2} \cdot \lambda_k^2 = \sum_{k=0}^{\infty} a_k^2 \lambda_k^{2m},$$

a convergent series. Hence Theorem 1 for $n = 2$, applied to $G_{m-1}(x)$, shows that $G'_{m-1}(x)$ is absolutely continuous and $\int_a^b [G''_{m-1}(x)]^2\, dx$ exists. As in (a)

$$(4.11)\qquad -rG_{m-1} = (pG'_{m-2})' - qG_{m-2} = \frac{p^{m-1}}{r^{m-2}} f^{(2m-2)}(x) + \mathfrak{L}(f, f', \cdots, f^{(2m-3)}).$$

Differentiation shows that $f^{(2m-1)}(x)$ is absolutely continuous, and a second differentiation shows that $f^{(2m)}(x)$ exists almost everywhere and that its square is integrable.

Proof of the corollary. The inequality (4.9) shows that the series

$$f^{(s)}(x) = \sum_{k=0}^{\infty} a_k \varphi_k^{(s)}(x) \qquad (s = 0, 1, 2, \cdots, n-1)$$

converge absolutely and uniformly, and the first part of the corollary is proved. The second part was proved for $n = 1$ under Case 1. If $n = 2m + 1$, (1) shows that

$$-r\lambda_k \varphi_k = (p\varphi_k')' - q\varphi_k .$$

Multiplication by $a_k \lambda_k^s$ gives

$$-ra_k\lambda_k^{s+1}\varphi_k = (pa_k\lambda_k^s\varphi_k')' - qa_k\lambda_k^s\varphi_k \qquad (s = 0, 1, 2, \cdots, m-1).$$

Since $a_k\lambda_k^s$ is the k-th Fourier coefficient of $G_s(x)$, this relation shows that each term of the series for $G_s(x)$ satisfies the same recursion relation that $G_s(x)$ satisfies. That is, if

$$A_k^{(s)} = a_k\lambda_k^s ,$$

then

$$-rA_k^{(s)}\varphi_k = (pA_k^{(s-1)}\varphi_k')' - qA_k^{(s-1)}\varphi_k \qquad (s = 1, 2, \cdots, m). \tag{4.12}$$

In particular,

$$\begin{aligned} -rA_k^{(1)}\varphi_k &= (pA_k^{(0)}\varphi_k')' - qA_k^{(0)}\varphi_k , \\ -rA_k^{(2)}\varphi_k &= (pA_k^{(1)}\varphi_k')' - qA_k^{(1)}\varphi_k . \end{aligned} \tag{4.13}$$

Differentiation of (4.13) gives

$$\begin{aligned} -rA_k^{(1)}\varphi_k' &= r'A_k^{(1)}\varphi_k + A_k^{(0)}(p\varphi_k')'' - A_k^{(0)}(q\varphi_k)' \\ &= -\frac{r'}{r}A_k^{(0)}(p\varphi_k')' + \frac{qr'}{r}A_k^{(0)}\varphi_k + A_k^{(0)}(p\varphi_k')'' - A_k^{(0)}(q\varphi_k)' \end{aligned}$$

from (4.13). That is, $A_k^{(1)}\varphi_k'$ is expressed as a linear combination of $A_k^{(0)}\varphi_k$, $A_k^{(0)}\varphi_k'$, $A_k^{(0)}\varphi_k''$ and $A_k^{(0)}\varphi_k'''$, the coefficient of the last being $-p/r$ and the other coefficients being polynomials in $p(x)$, $q(x)$, $1/r(x)$ and their derivatives of orders 2, 1 and 1 respectively. Writing (4.12) successively for $s = 1, 2, \cdots, m$ and repeating this process show that

$$-rA_k^{(m)}\varphi_k = -ra_k\lambda_k^m\varphi_k = \frac{p^m}{r^{m-1}}a_k\varphi_k^{(2m)} + \mathfrak{L}(a_k\varphi_k, a_k\varphi_k', \cdots, a_k\varphi_k^{(2m-1)}), \tag{4.14}$$

where $\mathfrak{L}(\qquad)$ is the same linear form as (4.10). Summation of (4.14) from $k = 0$ to $k = N$ and subtraction from (4.10) give

$$\begin{aligned} -r\Big(G_m - \sum_{k=0}^{N} a_k\lambda_k^m\varphi_k\Big) &= \frac{p^m}{r^{m-1}}\Big(f^{(2m)} - \sum_{k=0}^{N} a_k\varphi_k^{(2m)}\Big) \\ &\quad + \mathfrak{L}\Big(\Big[f - \sum_{k=0}^{N} a_k\varphi_k\Big], \Big[f' - \sum_{k=0}^{N} a_k\varphi_k'\Big], \cdots, \Big[f^{(2m-1)} - \sum_{k=0}^{N} a_k\varphi_k^{(2m-1)}\Big]\Big). \end{aligned}$$

Differentiation and rearrangement give

$$-r(G_m' - \sum_{k=0}^{N} a_k\lambda_k^m \varphi_k') - r'(G_m - \sum_{k=0}^{N} a_k\lambda_k^m \varphi_k)$$
$$+ \mathfrak{L}_1([f - \sum_{k=0}^{N} a_k\varphi_k], \cdots, [f^{(2m)} - \sum_{k=0}^{N} a_k\varphi_k^{(2m)}])$$
$$= \frac{p^m}{r^{m-1}}(f^{(2m+1)} - \sum_{k=0}^{N} a_k\varphi_k^{(2m+1)}),$$

where $\mathfrak{L}_1(\qquad)$ is a linear form whose coefficients are continuous. Hence

$$\int_a^b \frac{p^{2m}}{r^{2m-2}}(f^{(2m+1)} - \sum_{k=0}^{N} a_k\varphi_k^{(2m+1)})^2\,dx \leqq 3\int_a^b r^2(G_m' - \sum_{k=0}^{N} a_k\lambda_k^m\varphi_k')^2\,dx$$
$$+ 3\int_a^b (r')^2(G_m - \sum_{k=0}^{N} a_k\lambda_k^m\varphi_k)^2\,dx + 3\int_a^b \mathfrak{L}_1^2\,dx.$$

$G_m(x)$ satisfies the hypotheses of the corollary for $n = 1$, as was shown in the second paragraph of Case 3. Consequently the first integral on the right approaches zero as $N \to \infty$. In the other integrals on the right each series converges uniformly to the function from which it is being subtracted. Hence these integrals approach zero as $N \to \infty$, and the corollary is proved.

If $n = 2m$, it was shown in the proof of Case 3 (b) that Theorem 1 for $n = 2$ could be applied to $G_{m-1}(x)$ to show that $G_{m-1}'(x)$ is absolutely continuous and $\int_a^b [G_{m-1}''(x)]^2\,dx$ exists. The Fourier coefficients of $G_{m-1}(x)$ are $a_k\lambda_k^{m-1}$, as was shown in Lemma 5. Hence

$$\lim_{N\to\infty}\int_a^b (G_{m-1}'' - \sum_{k=0}^{N} a_k\lambda_k^{m-1}\varphi_k'')^2\,dx = 0,$$

and (4.11) proves the corollary, by an argument similar to the preceding.

5. **Proof of Theorem 2.**

Case 1: $n = 1$. *If $f(x)$ is an absolutely continuous function such that $\int_a^b (f'(x))^2\,dx$ exists and $f(x)$ satisfies any boundary conditions $y(a) = 0$, $y(b) = 0$, $y(a) = Ky(b)$ which may be implied by* (2), *then $\sum_{k=0}^{\infty} a_k^2\lambda_k$ converges.*

Proof. If $F(x)$ satisfies the hypotheses of the theorem, and if in addition $F''(x)$ is piecewise continuous ($a \leqq x \leqq b$), let $A_k = \int_a^b Fr\varphi_k\,dx$. Let

$$\int_a^b pF''\varphi_k\,dx + \int_a^b p'F'\varphi_k\,dx = b_k + c_k.$$

Integration by parts gives

$$
\begin{aligned}
b_k + c_k &= F' p\varphi_k \Big|_a^b - \int_a^b F'(p\varphi_k' + p'\varphi_k)\,dx + \int_a^b p'F'\varphi_k\,dx \\
&= p(F'\varphi_k - F\varphi_k')\Big|_a^b + \int_a^b F(q - \lambda_k r)\varphi_k\,dx
\end{aligned}
\tag{5.1}
$$

after a second integration by parts and the use of (1). In the three cases of the boundary conditions listed at the end of Lemma 4 the integrated terms become respectively

$$A\varphi_k(b) + B\varphi_k(a), \qquad A\varphi_k(b), \qquad B\varphi_k(a) \text{ or zero.}$$

The arguments are those used in the simplification of (4.3), if it is remembered that $F(x)$ satisfies some of the boundary conditions if (3.5) or (3.6) holds. Then if $d_k = \int_a^b Fq\varphi_k\,dx$, (5.1) gives

$$
\begin{aligned}
A_k\lambda_k &= A\varphi_k(b) + B\varphi_k(a) - b_k - c_k + d_k, \\
A_k^2\lambda_k^2 &\leqq 5[A^2\varphi_k^2(b) + B^2\varphi_k^2(a) + b_k^2 + c_k^2 + d_k^2], \\
A_k^2\lambda_k &\leqq 5\left[A^2\frac{\varphi_k^2(b)}{\lambda_k} + B^2\frac{\varphi_k^2(a)}{\lambda_k} + \frac{b_k^2}{\lambda_k} + \frac{c_k^2}{\lambda_k} + \frac{d_k^2}{\lambda_k}\right].
\end{aligned}
\tag{5.2}
$$

Here A and B are constants depending on $p(a)$, $p(b)$, $F(a)$, $F(b)$, $F'(a)$ and $F'(b)$ only, and b_k, c_k and d_k are respectively the k-th Fourier coefficients of the piecewise continuous functions pF''/r, $p'F'/r$, and Fq/r. Therefore $\sum_{k=0}^{\infty} b_k^2$, $\sum_{k=0}^{\infty} c_k^2$ and $\sum_{k=0}^{\infty} d_k^2$ converge, and Lemma 3 and (5.2) show then that $\sum_{k=0}^{\infty} A_k^2\lambda_k$ converges. Hence $\sum_{k=0}^{\infty} A_k\varphi_k$ converges absolutely and uniformly to $F(x)$, by the argument used in Theorem 1, Case 1 and the remark at the end of §2.

To complete the proof the three cases of the boundary conditions will be dealt with separately. If (3.4) holds, let $f(x)$ be a function satisfying the hypotheses of the theorem and let $F(x)$ be a function with piecewise continuous second derivative such that

$$\int_a^b (F' - f')^2\,dx < \frac{\epsilon^2}{b - a}, \tag{5.3}$$

where ϵ will be specified later. Then

$$| F(x) - f(x) | \leqq | F(\xi) - f(\xi) | + \left| \int_\xi^x (F' - f')\,dx \right|,$$

where ξ is any number in the closed interval (a, b). Therefore

$$| F(x) - f(x) | \leqq | F(\xi) - f(\xi) | + \left| \int_\xi^x (F' - f')^2\,dx \cdot (x - \xi) \right|^{\frac{1}{2}}$$

by Schwarz' inequality; a fortiori

$$(5.4)\qquad | F(x) - f(x) | \leqq | F(\xi) - f(\xi) | + \left[\int_a^b (F' - f')^2\, dx \cdot (b - a)\right]^{\frac{1}{2}}.$$

If ξ is fixed arbitrarily and $F(\xi)$ is chosen equal to $f(\xi)$,

$$(5.5)\qquad | F(x) - f(x) | < \epsilon \qquad (a \leqq x \leqq b).$$

Now

$$0 \leqq \int_a^b p(F' - \sum_{k=0}^N A_k\varphi_k')^2\, dx$$

$$= \int_a^b pF'^2\, dx - 2\sum_{k=0}^N A_k \int_a^b F'p\varphi_k'\, dx + \sum_{k=0}^N \sum_{l=0}^N A_k A_l \int_a^b p\varphi_k'\varphi_l'\, dx.$$

Integration of the second and third terms by parts and use of (1) give

$$0 \leqq \int_a^b pF'^2\, dx - 2\sum_{k=0}^N A_k \left[pF\varphi_k' \Big|_a^b - \int_a^b F(q - \lambda_k r)\varphi_k\, dx\right]$$

$$+ \sum_{k=0}^N \sum_{l=0}^N A_k A_l \left[p\varphi_k'\varphi_l \Big|_a^b - \int_a^b \varphi_l(q - \lambda_k r)\varphi_k\, dx\right]$$

$$(5.6)\qquad = \int_a^b pF'^2\, dx + 2\sum_{k=0}^N A_k \int_a^b Fq\varphi_k\, dx - \sum_{k=0}^N \sum_{l=0}^N A_k A_l \int_a^b q\varphi_k\varphi_l\, dx$$

$$- \sum_{k=0}^N A_k^2\lambda_k - 2\sum_{k=0}^N A_k pF\varphi_k' \Big|_a^b + \sum_{k=0}^N \sum_{l=0}^N A_k A_l p\varphi_k'\varphi_l \Big|_a^b.$$

By the use of (3.4) the integrated terms can be written

$$Q = A\sum_{k=0}^N A_k\varphi_k(b) + B\sum_{k=0}^N A_k\varphi_k(a) + C(\sum_{k=0}^N A_k\varphi_k(b))^2$$

$$+ D(\sum_{k=0}^N A_k\varphi_k(a))^2 + E(\sum_{k=0}^N A_k\varphi_k(a))(\sum_{k=0}^N A_k\varphi_k(b)).$$

The series are all uniformly convergent to $F(a)$ or $F(b)$, as was shown in the first part of the proof, and A, B, C, D and E are constants depending only on $F(a)$, $F(b)$, $p(a)$, $p(b)$, $1/\delta_{34}$ and the δ_{ij}. Then

$$\sum_{k=0}^N A_k^2\lambda_k \leqq \int_a^b pF'^2\, dx + 2\int_a^b Fq(\sum_{k=0}^N A_k\varphi_k)\, dx - \int_a^b q(\sum_{k=0}^N A_k\varphi_k)^2\, dx + Q.$$

$\lim_{N\to\infty} Q$ is the sum of a finite number of products of some of $F(a)$, $F(b)$, $p(a)$, $p(b)$, $1/\delta_{34}$ and the δ_{ij}. Then (5.5) shows that $\lim_{N\to\infty} Q < M$, a constant involving similar products of some of $f(a)$, $f(b)$, $p(a)$, $p(b)$, $1/\delta_{34}$ and the δ_{ij}. Hence

$$(5.7)\qquad \sum_{k=0}^N A_k^2\lambda_k \leqq \sum_{k=0}^\infty A_k^2\lambda_k \leqq \int_a^b pF'^2\, dx + \int_a^b qF^2\, dx + M,$$

where M is independent of $F(x)$.

Now let ϵ and $F(x)$ be chosen so that

$$\left| \int_a^b pF'^2 \, dx - \int_a^b pf'^2 \, dx \right| < 1 \qquad \text{and} \qquad \left| \int_a^b qF^2 \, dx - \int_a^b qf^2 \, dx \right| < 1.$$

The first inequality can be obtained because of (5.3) and the second because of (5.5). Then (5.7) can be replaced by

$$(5.8) \qquad \sum_{k=0}^{N} A_k^2 \lambda_k \leqq \sum_{k=0}^{\infty} A_k^2 \lambda_k \leqq \int_a^b pf'^2 \, dx + \int_a^b qf^2 \, dx + M + 2.$$

For every fixed N, $F(x)$ can be still further restricted so that

$$\left| \sum_{k=0}^{N} A_k^2 \lambda_k - \sum_{k=0}^{N} a_k^2 \lambda_k \right| < 1,$$

since by (5.5)

$$| A_k - a_k | = \left| \int_a^b (F - f) r \varphi_k \, dx \right| < \epsilon \int_a^b r \, | \varphi_k | \, dx \qquad (k = 0, 1, \cdots, N).$$

Then (5.8) can be replaced by

$$\sum_{k=0}^{N} a_k^2 \lambda_k \leqq \int_a^b pf'^2 \, dx + \int_a^b qf^2 \, dx + M + 3.$$

This inequality holds for every N and its right member is independent of $F(x)$. Therefore $\sum_{k=0}^{\infty} a_k^2 \lambda_k$ converges.

If (3.6) holds, the proof is similar. The point $x = \xi$ in (5.4) is chosen as $x = b$. The integrated terms in (5.6) vanish at $x = b$ and at the other end-point $h_2 \varphi_k'(a) = h_1 \varphi_k(a)$.[14]

In the third case (3.5) holds. Let $G(x)$ be a function of class C″ such that

$$\int_a^b (G' - f')^2 \, dx < \frac{\delta^2}{b - a}, \qquad \text{where } \delta < (b - a)^2,$$

and $G(a) = f(a)$. Then

$$(5.9) \qquad | G(x) - f(x) | < \delta \qquad (a \leqq x \leqq b),$$

as before. Now let

$$F(x) = \begin{cases} G(x), & a \leqq x \leqq b - \delta^{\frac{1}{2}}, \\ G(x) + C_0 x^2 + C_1 x + C_2, & b - \delta^{\frac{1}{2}} \leqq x \leqq b. \end{cases}$$

[14] The result in this case can also be proved by using $f(x)$ instead of $F(x)$ in (5.6). The problem is a Sturm-Liouville problem and it is known that $f(x) = \sum_{k=0}^{\infty} a_k \varphi_k(x)$. This makes the approximation by $F(x)$ unnecessary.

The constants C_i are determined by the conditions

$$F(b - \delta^{\frac{1}{2}}) = G(b - \delta^{\frac{1}{2}}), \qquad F'(b - \delta^{\frac{1}{2}}) = G'(b - \delta^{\frac{1}{2}}), \qquad F(b) = f(b).$$

This means

$$C_0 = \frac{f(b) - G(b)}{\delta}, \qquad C_1 = -2(b - \delta^{\frac{1}{2}})\frac{f(b) - G(b)}{\delta},$$

$$C_2 = f(b) - G(b) - C_0 b^2 - C_1 b.$$

By (5.9) $| C_0 | < 1$ and C_1 and C_2 are bounded. Then

$$\int_{b-\delta^{\frac{1}{2}}}^{b} F'^2\, dx \leqq 2\int_{b-\delta^{\frac{1}{2}}}^{b} G'^2\, dx + 2\int_{b-\delta^{\frac{1}{2}}}^{b} (2C_0 x + C_1)^2\, dx = 2\int_{b-\delta^{\frac{1}{2}}}^{b} G'^2\, dx + O(\delta^{\frac{1}{2}}).$$

Thus $F(x)$ is a function whose second derivative is piecewise continuous ($a \leqq x \leqq b$), the only possible discontinuity being a jump at $x = b - \delta^{\frac{1}{2}}$. Furthermore $F(x)$ satisfies the same first boundary condition $\delta_{14}F(a) = -\delta_{24}F(b)$ as $f(x)$, and given any $\epsilon > 0$

$$\int_a^b [F'(x) - f'(x)]^2\, dx < \epsilon$$

if δ is suitably chosen. Then the previous proof can be carried through.

Case 2: $n = 2$. *The hypotheses are that $f'(x)$ is absolutely continuous, $f(x)$ satisfies the boundary conditions, and $\int_a^b [f''(x)]^2\, dx$ exists.* The equation (5.1) is now

$$b_k + c_k = p(f'\varphi_k - f\varphi_k')\Big|_a^b + \int_a^b fq\varphi_k\, dx - \lambda_k \int_a^b fr\varphi_k\, dx.$$

The integrated terms all vanish because $f(x)$ satisfies the boundary conditions, as is seen by substitution. The three cases of the boundary conditions can be considered separately. Then (5.2) becomes merely

$$a_k^2\lambda_k^2 \leqq 3[b_k^2 + c_k^2 + d_k^2],$$

so that $\sum_{k=0}^{\infty} a_k^2\lambda_k^2$ converges.

Case 3: $n > 2$. The proof is by induction. If $n = 1$ or 2, the theorem has already been proved. If $n = 2m$, the theorem is assumed true for $n = 2m - 1$ and the hypotheses for $n = 2m$ are assumed to be satisfied; viz., *$f^{(2m-1)}(x)$ is absolutely continuous ($a \leqq x \leqq b$), $\int_a^b [f^{(2m)}(x)]^2\, dx$ exists, and $G_0 \equiv f, G_1, \cdots, G_{m-1}$ satisfy the boundary conditions. The conclusion is that $\sum_{k=0}^{\infty} a_k^2\lambda_k^{2m}$ converges.*

From the definition

$$-rG_{m-1} = (pG_{m-2}')' - qG_{m-2} = \mathfrak{L}(f, f', \cdots, f^{(2m-2)}),$$

a linear combination of the indicated derivatives of $f(x)$, the coefficients being polynomials in $p(x)$, $q(x)$ and $1/r(x)$ and their derivatives of orders not greater than $2m - 3$, $2m - 4$ and $2m - 4$, respectively. The hypotheses on $f(x), p(x)$, $q(x)$ and $r(x)$ imply that $d\mathfrak{L}/dx$ is absolutely continuous, that $d^2\mathfrak{L}/dx^2$ exists almost everywhere and that $\int_a^b [d^2\mathfrak{L}/dx^2]^2\,dx$ exists. Since $r(x)$ has at least two continuous derivatives and is nowhere zero, it follows that $G'_{m-1}(x)$ is absolutely continuous and $\int_a^b [G''_{m-1}(x)]^2\,dx$ exists. Consequently $G_{m-1}(x)$ satisfies the hypotheses of the theorem for $n = 2$. Then if

$$A_k = \int_a^b G_{m-1} r\varphi_k\,dx,$$

$\sum_{k=0}^{\infty} A_k^2\lambda_k^2$ converges.

The hypotheses of the theorem for $n = 2m$ are sufficient for $n = 2m - 1$, and the theorem is therefore true by the induction hypothesis. Then the argument of Theorem 1, Case 2 shows that $f(x) \equiv \sum_{k=0}^{\infty} a_k\varphi_k(x)$, since the set of functions $\{\varphi_k(x)\}$ is known to be complete. Then Lemma 5 shows that

$$G_{m-1}(x) = \sum_{k=0}^{\infty} a_k\lambda_k^{m-1}\varphi_k(x),$$

so that

$$A_k = a_k\lambda_k^{m-1}.$$

Hence $\sum_{k=0}^{\infty} a_k^2\lambda_k^{2m} = \sum_{k=0}^{\infty} A_k^2\lambda_k^2$ converges.

If $n = 2m + 1$, the theorem is assumed true for $n = 2m$ and the further hypotheses are now: *$f^{(2m)}(x)$ is absolutely continuous, $\int_a^b [f^{(2m+1)}(x)]^2\,dx$ exists, $G_0 \equiv f, G_1, \cdots, G_{m-1}$ satisfy the boundary conditions and $G_m(x)$ satisfies any boundary conditions $y(a) = 0$, $y(b) = 0$, $y(a) = Ky(b)$ which may be implied by* (2). Now

$$-rG_m = (pG'_{m-1})' - qG_{m-1}.$$

As in the previous case it follows that $G_m(x)$ is absolutely continuous and $\int_a^b [G'_m(x)]^2\,dx$ exists. Then if $B_k = \int_a^b G_m r\varphi_k\,dx$, the theorem for $n = 1$, applied to $G_m(x)$, shows that $\sum_{k=0}^{\infty} B_k^2\lambda_k$ converges.

But

$$B_k = -\int_a^b (pG'_{m-1})'\varphi_k\,dx + \int_a^b qG_{m-1}\varphi_k\,dx$$

$$= -pG'_{m-1}\varphi_k\Big|_a^b + \int_a^b pG'_{m-1}\varphi'_k\,dx + \int_a^b qG_{m-1}\varphi_k\,dx,$$

or

$$(5.10)\qquad B_k = -[p(G'_{m-1}\varphi_k - G_{m-1}\varphi'_k)]_a^b - \int_a^b G_{m-1}(p\varphi'_k)'\,dx + \int_a^b qG_{m-1}\varphi_k\,dx.$$

The integrated terms vanish on account of the boundary conditions, and use of (1) reduces (5.10) to

$$B_k = \lambda_k \int_a^b G_{m-1} r\varphi_k\,dx.$$

As before Lemma 5 shows the right member to be

$$\lambda_k a_k \lambda_k^{m-1}.$$

Hence $\sum_{k=0}^{\infty} a_k^2 \lambda_k^{2m+1} = \sum_{k=0}^{\infty} B_k^2 \lambda_k$ converges.[15]

THE UNIVERSITY OF ROCHESTER AND BROWN UNIVERSITY.

[15] Since

$$\Big(\sum_{k=N}^{N+M} |a_k|\,\lambda_k^t\,|\varphi_k^{(j)}|\Big)^2 = \left(\sum_{k=N}^{N+M} |a_k|\,\lambda_k^{t+\frac{1}{2}j+\frac{1}{2}} \cdot \frac{|\varphi_k^{(j)}|}{\lambda_k^{\frac{1}{2}j+\frac{1}{2}}}\right)^2 \qquad (j = 0 \text{ or } 1)$$

$$\leqq \Big(\sum_{k=N}^{N+M} a_k^2 \lambda_k^{2t+j+1}\Big) \cdot \left(\sum_{k=N}^{N+M} \frac{(\varphi_k^{(j)})^2}{\lambda_k^{j+1}}\right),$$

the argument from Lemma 3 used before shows that the series $\sum_{k=0}^{\infty} a_k \lambda_k^t \varphi_k^{(j)}(x)$ are absolutely and uniformly convergent if $2t + j + 1 \leqq n$. Hence the series for $G_0(x), G_1(x), \cdots, G_{m-1}(x)$, $G'_0(x), G'_1(x), \cdots, G'_{m-1}(x)$ (and $G_m(x)$ if n is odd) are absolutely and uniformly convergent, and since they satisfy the boundary conditions term by term, *the hypotheses* (b) *of Theorem* 2 *are seen to be necessary conditions*.

THE JACOBI CONDITION FOR UNILATERAL VARIATIONS

By J. D. Mancill

The simplest type of parametric problem of the calculus of variations is that of minimizing an integral

$$J = \int_{t_1}^{t_2} F(x, y, x', y')\, dt$$

in a class of admissible curves

$$(1) \qquad x = x(t), \qquad y = y(t) \qquad (t_1 \leqq t \leqq t_2)$$

joining two fixed points 1 and 2 in the xy-plane.[1] We shall suppose that the integrand function $F(x, y, x', y')$ satisfies the usual continuity and homogeneity properties for (x, y) in a region R of the plane and for all $(x', y') \neq (0, 0)$.[2] A curve (1) will be called *admissible* if it lies in or on the boundary of a region R' within the region R and is of class[3] C' except possibly at a finite number of corners.

A familiar formulation of the Jacobi condition states that there shall be no pair of conjugate points on any arc of a minimizing curve E_{12} which lies interior to the region R', is of class C', and has $F_1 \neq 0$ along it. The well-known geometric proof of this condition by means of the Kneser envelope theorem of a one-parameter family of extremals does not necessarily apply to every extremal arc of E_{12} in the case when E_{12} may have arcs in common with the boundary of the region R' of admissible variations, and so far as the author knows, no use has yet been made of the second variation in this connection. There is no restriction on the length of the arcs of the minimizing curve in common with the boundary of R' which are not extremals as the sufficient conditions of Bliss[4] show. The purpose of this paper is to show that by means of the Kneser envelope theorem and two additional properties of the envelope we are able to complete the proof of the Jacobi condition for the situation just described.[5]

Suppose, for sake of illustration, that the minimizing curve E_{12} is an *extremal*

Received July 31, 1939; presented to the American Mathematical Society, December 29, 1939.

[1] We shall denote all derivatives with respect to the independent variable t by primes and all others by subscripts.

[2] See, e.g., O. Bolza, *Lectures on the Calculus of Variations*, University of Chicago, p. 117.

[3] For a definition of the term class as here used see Bolza, loc. cit., p. 116.

[4] G. A. Bliss, *Sufficient conditions for a minimum with respect to one-sided variations*, Transactions of the American Mathematical Society, vol. 5(1904), p. 491.

[5] Incidentally, the results of this paper are of historical interest since they eliminate the restriction, in the case of the geometric formulation, that the envelope have a regressive branch at its points of contact with the minimizing curve.

and contains an arc, say E_{34}, in common with the boundary B of the region R'. Suppose also that the function $F_1(x, y, x', y')$ is different from zero along E_{12}. Let

$$(2) \qquad x = \phi(t, a), \qquad y = \psi(t, a)$$

represent the one-parameter family of extremals through the point 1 which contains E_{12} for $t_1 \leqq t \leqq t_2$, $a = a_0$. We shall now prove the

THEOREM. *If the extremal E_{12} contains a point of contact with the envelope of the family of extremals* (2), *excepting possibly the point* 2, *then E_{12} cannot minimize the integral J.*

Suppose that E_{12} has the point 0 as its *first* point of contact with the envelope of the family of extremals (2). Then the equation

$$D(t, a) = \begin{vmatrix} \phi_t & \phi_a \\ \psi_t & \psi_a \end{vmatrix} = 0$$

has the solution $(t, a) = (t_0, a_0)$ and thus determines the function $t = t(a)$ which is of class C' near $a = a_0$, since the derivative $D_t(t_0, a_0)$ is not zero.[6] Consequently, the equations of the enveloping curve of the family (2) takes the form

$$e: \qquad X(a) \equiv \phi[t(a), a], \qquad Y(a) \equiv \psi[t(a), a],$$

where the functions $X(a)$ and $Y(a)$ are of class C' near $a = a_0$.

Let us suppose that the functions $X(a)$ and $Y(a)$ are of class $C^{(r)}$ in the vicinity of $a = a_0$,[7] that for $a = a_0$ their derivatives of order not exceeding $r - 1$ vanish, but that the r-th derivatives do not both vanish. Then Bolza[8] has shown that the relations

$$\frac{dX}{d\tau} = m\phi_t, \qquad \frac{dY}{d\tau} = m\psi_t,$$

where $a - a_0 = \epsilon\tau$, $\epsilon = \pm 1$, $m = \epsilon^r\tau^{r-2}(n + \nu)$, $n \neq 0$, and $\lim_{\tau=0} \nu = 0$, hold near $\tau = 0$. An easy calculation shows that the relation

$$D_t[t(\tau), \tau](x'^2 + y'^2)^{-\frac{1}{2}} = m\left(\frac{1}{\rho} - \frac{1}{r}\right),$$

where $1/\rho$ and $1/r$ are the curvatures of e and that member of the family (2) tangent to e respectively, holds in the vicinity of $\tau = 0$. From this relation it is easily seen that *in the subcase* (A) *considered by Bolza,*[9] *the curve e lies entirely on one side of E_{12} in the vicinity of the point* 0 *and that in his subcase* (B) *the curve e crosses the extremal E_{12} at the point* 0.

[6] Bolza, loc. cit., p. 200.

[7] This is the case, for example, when the integrand function F is of class $C^{(r+3)}$.

[8] Loc. cit., pp. 201–202.

[9] Loc. cit., p. 202.

The theorem can now be proved by means of the well-known Kneser envelope theorem in the subcase (B_1) of Bolza and also in his subcase (A) if the envelope e lies entirely on the left (when the admissible region lies to the left of the positive direction of the boundary B) of E_{12} near the point 0. We shall suppose that the parameter a is chosen so that the branch of e under consideration and the extremal E_{12} have the same positive direction at the point 0. Now, in subcase (B_2) of Bolza and in his subcase (A) when e lies to the right of E_{12} in the vicinity of the point 0, consider the members of the family (2) tangent to e for $a_0 \leqq a \leqq a_0 + \epsilon$. These members intersect E_{12} beyond the point 0 for ϵ sufficiently small, since the branch of e under consideration is entirely on the right of E_{12} near a_0 and these members are tangent to e beyond a_0. We shall now show that they simply cover the region of the xy-plane determined by the equations (2) for

$$R_{ta}: \qquad a_0 \leqq a \leqq a_0 + \epsilon, \qquad t(a) - \delta \leqq t \leqq t(a),$$

if ϵ and δ are sufficiently small.

We may suppose that $\phi'(t_0, a_0) > 0$, since E_{12} is non-singular at 0, and thus every ordinate in the interval $x_0 \leqq x \leqq x_0 + \beta$ defined by R_{ta} is crossed once and only once by each extremal arc defined in this range if ϵ and δ are sufficiently small, since E_{12} crosses the ordinate $x = x_0$. Consider the unique value of t for which a given extremal corresponding to a value of a crosses the ordinate x. This is a function $t(x, a)$ which satisfies the equation

$$x = \phi(t, a)$$

and by the usual theorems on implicit functions is therefore of class C′ near (x_0, a_0). The only points (x, y) in the interval $x_0 \leqq x \leqq x_0 + \beta$ corresponding to the region R_{ta} by means of the equations (2) are those which satisfy the equation

$$y = \psi[t(x, a), a] = y(x, a). \tag{3}$$

The derivative of $y(x, a)$ with respect to a is readily found to be

$$y_a(x, a) = \frac{D[t(x, a), a]}{\phi'[t(x, a), a]}.$$

Therefore, on any ordinate $x_0 \leqq x \leqq x_0 + \beta$ the value of y continuously increases or decreases for a in R_{ta} according as the sign of the determinant D is positive or negative, since D retains its sign for the range R_{ta}. It is thus clear that the points of the region R_{ta} are in one-to-one correspondence with those of the region defined by the equations (2) if ϵ and δ are sufficiently small. The single-valued functions $t = t(x, y)$ and $a = a(x, y)$ which determine the unique extremal of the family (2) through each point of this region, are of class C′ interior to the region, since for such values (t, a) the functional determinant $D(t, a)$ of equations (2) is different from zero, and are continuous on the boundary.

Let us consider the family of admissible curves

$$E_u: \quad \begin{array}{lll} x = \phi[t, a(u)], & y = \psi[t, a(u)], & t_1 \leqq t \leqq t(u), \\ x = \bar{x}(t) \equiv \phi(t, a_0), & y = \bar{y}(t) \equiv \psi(t, a_0), & u \leqq t \leqq t_2, \end{array}$$

where $t(u) = t[x(u), y(u)]$ and $a(u) = a[x(u), y(u)]$ determine the unique extremal of the family (2) through each point of the curve E_{12} in the interval $u_0 \leqq u \leqq u_0 + \epsilon$. The function

$$I(u) = I(E_u)$$

is continuous on $u_0 \leqq u \leqq u_0 + \epsilon$ and has a derivative on $u_0 < u \leqq u_0 + \epsilon$ for ϵ sufficiently small. By making use of the extremal property of the members of the family (2) and the homogeneity property of the integrand function F, it can be shown that the derivative of $I(u)$ takes the form

$$(4) \qquad \begin{aligned} I_u(u) &= F_{x'}\bar{x}' + F_{y'}\bar{y}' \,\Big|^{t(u)} - F(x, y, x', y')\,\Big|^{u} \\ &= -\mathcal{E}(x, y, x', y', \bar{x}', \bar{y}')\,\Big|^{t(u)}, \end{aligned}$$

where $\mathcal{E}(x, y, x', y', \bar{x}', \bar{y}')$ is the Weierstrass $\mathcal{E}$-function. Since the curve E_{12} is supposed to minimize the integral I, it follows that $\mathcal{E}(x, y, x', y', p, q) \geqq 0$ along E_{12} for all $(p, q) \neq (0, 0)$.[10] If the function F_1 is different from zero along E_{12}, then $F_1(x, y, x', y') > 0$ for all elements (x, y, x', y') in a certain neighborhood of those along E_{12}. Hence[11] $\mathcal{E}(x, y, x', y', p, q) > 0$ along each member of the family (2) sufficiently near E_{12} for all directions (p, q) sufficiently near the direction (x', y') of E_{12}. Also, we have

$$\lim_{u=u_0} [\phi'(t(u), a(u)), \psi'(t(u), a(u))] = (\bar{x}', \bar{y}').$$

Therefore, it follows that

$$\mathcal{E}[\phi(t, a(u)), \psi(t, a(u)), \phi'(t, a(u)), \psi'(t, a(u)), \bar{x}', \bar{y}']^{t(u)} > 0$$

for $u_0 < u < u_0 + \epsilon$ if ϵ is sufficiently small. Thus it follows from equations (4) that $I_u(u) < 0$ for $u > u_0$ but sufficiently near u_0. Therefore, the curve E_{12} could not minimize the integral I.

UNIVERSITY OF ALABAMA.

[10] L. M. Graves, *A proof of the Weierstrass condition in the calculus of variations*, American Mathematical Monthly, vol. 41(1934), pp. 502–504.

[11] Cf. Bolza, loc. cit., p. 140.

DIVISIBILITY PROPERTIES OF INTEGRAL FUNCTIONS

By Olaf Helmer

1. **Introduction.** So much is known today about the divisibility properties of polynomials and the relation between their zeros and their coefficients, and so little about the corresponding properties of integral functions, that it is tempting to try to transfer some of the results of the arithmetic and algebra of polynomials to the wider field of integral functions. This paper is meant to represent an initial step in this direction.

The step from a finite to an infinite number of coefficients and zeros necessitates, of course, a change of method which, naturally, is a change for the worse, since the most efficient methods do not apply any longer. For example, the powerful device of complete induction with respect to the degree is lost.

This paper is devoted to a first survey of the divisibility relations among integral functions. The situation resembles to some extent the earliest stage of the divisibility theory of algebraic numbers. It is likewise characterized by the presence of a disturbing abundance of divisors. This remains so, even when the coefficients are restricted to a subfield of the field of all complex numbers, so that some further restriction seems to be indicated. Later, at least one very essential reason will become apparent for carrying this restriction into effect by confining the study of the divisibility structure to domains of integral functions of *finite* order with coefficients in a given field.

2. **Notation.** The fields of rational, real, or complex numbers may be denoted by P, R, or C, respectively. As usual, $K[z]$ stands for the domain of polynomials in z with coefficients in K, where K is an arbitrary subfield of C. By $K\langle z\rangle$ let the domain of integral functions in z with coefficients in K be understood, that is, the set of all power series $f(z) = \sum a_n z^n$ with a_n in K and $\lim |a_n|^{1/n} = 0$.

The order of growth of a function $f(z)$, written ord f, may be finite or infinite. The subset of those $f(z)$ of $K\langle z\rangle$ with ord $f < \infty$ will be denoted by $K^*\langle z\rangle$.

Let Z_f be the set of all zeros of $f(z)$, each occurring a number of times equal to its multiplicity; thus Z_f is an "algebraic" set, in the sense of possibly containing some of its elements several times. Z_f may be called *symmetric* if its elements are distributed symmetrically with respect to the real axis (so that two conjugate complex numbers occur with the same multiplicity). An algebraic set (in the above sense) may be called a *zero-set,* if it has no finite accumulation point. Z is, of course, a zero-set if and only if there is an integral function $f(z)$ such that $Z = Z_f$.

Received August 19, 1939.

3. **The units in $K\langle z\rangle$ and $K^*\langle z\rangle$.** Both $K\langle z\rangle$ and $K^*\langle z\rangle$ are obviously integral domains. Their units, that is, their divisors of unity, can easily be determined.

THEOREM 1. *An element of $K\langle z\rangle$ is a unit of $K\langle z\rangle$ if and only if it has the form $c\cdot e^{f(z)}$ with $c \neq 0$, in K, and $f(z)$ in $K\langle z\rangle$, $f(0) = 0$.*

Proof. A unit of $K\langle z\rangle$ has no zeros and must therefore be of the form $e^{g(z)}$, or $c\cdot e^{f(z)}$, where $c = e^{g(0)}$ and $f(0) = 0$. On the other hand, any such function *is* a unit because of $e^{g}\cdot e^{-g} = 1$, provided its coefficients are in K. For the latter to be true it is necessary and sufficient that c and the coefficients of

$$(1) \qquad e^{f(z)} = 1 + \frac{a_1 z + a_2 z^2 + \cdots}{1!} + \frac{(a_1 z + a_2 z^2 + \cdots)^2}{2!} + \cdots,$$

that is,

$$(2) \qquad a_1\,,\ a_2 + \tfrac{1}{2}a_1^2\,,\ a_3 + a_1 a_2 + \tfrac{1}{6}a_1^3\,,\ \cdots,$$

be in K. Now, evidently, the quantities (2) belong to K if and only if a_1, a_2, a_3, $\cdots$ are in K.

THEOREM 2. *An element of $K^*\langle z\rangle$ is a unit of $K^*\langle z\rangle$ if and only if it has the form $c\cdot e^{F(z)}$ with $c \neq 0$, in K, and $F(z)$ in $K[z]$, $F(0) = 0$.*

Proof. A unit of $K^*\langle z\rangle$ is also a unit of $K\langle z\rangle$ and hence has the form $c\cdot e^{F(z)}$; this function is of finite order if and only if $F(z)$ is a polynomial.

4. **Functions with given zero-sets.** If Z is any zero-set, we can, by virtue of Weierstrass' theorem, construct a function $f(z)$ in $C\langle z\rangle$ such that $Z = Z_f$. Moreover, if Z is a symmetric zero-set, such a function can be found in $R\langle z\rangle$. The following theorem shows that, in addition, a very strong restriction may be imposed upon the coefficients. This is a first indication of how decidedly the situation here differs from that in the case of polynomials.

THEOREM 3. *Let K be any imaginary number field, for example, $P(i)$; then, for every zero-set Z, there is a function $f(z)$ in $K\langle z\rangle$ such that $Z = Z_f$. Moreover, if Z is symmetric, such a function can already be found in $P\langle z\rangle$.*

Proof. To prove Theorem 3, we have to show that every integral function $f(z)$ in $R\langle z\rangle$ or in $C\langle z\rangle$ can be multiplied by a unit $e^{\varphi(z)}$ of such a kind that $g(z) = e^{\varphi(z)}\cdot f(z)$ has coefficients in P or in K, respectively, where K is a given imaginary field. Let

$$(3) \qquad f(z) = \sum a_n z^n, \qquad g(z) = \sum b_n z^n, \qquad \varphi(z) = \sum \lambda_n z^n.$$

We may assume $f(0) = a_0 \neq 0$ (otherwise, let $f(z) = z^s\cdot f_1(z)$ with $f_1(0) \neq 0$, $g_1(z) = e^{\varphi(z)}\cdot f_1(z)$, and $g(z) = z^s\cdot g_1(z)$). Moreover, since multiplication by a constant does not alter the zeros, we may assume $a_0 = 1$. Further, let $\lambda_0 = 0$ and hence $b_0 = 1$. We now have

$$(4) \quad \exp\,(\lambda_1 z + \lambda_2 z^2 + \cdots)(1 + a_1 z + a_2 z^2 + \cdots) = 1 + b_1 z + b_2 z^2 + \cdots.$$

Comparing coefficients of z^n in (4), we obtain equations of the form

$$P_n(a_1, a_2, \cdots, a_n, \lambda_1, \lambda_2, \cdots, \lambda_{n-1}) + \lambda_n = b_n, \tag{5}$$

where P_n is a polynomial in the quantities indicated, with certain rational coefficients. Our task is to determine the λ_n in such a manner that

$$\begin{cases} \text{(i) } \varphi(z) \text{ is an integral function,} \\ \text{(ii) } b_n \text{ in } P, \text{ or in } K, \text{ respectively.} \end{cases} \tag{6}$$

Now (i) is fulfilled if and only if $\lim |\lambda_n|^{1/n} = 0$. This will certainly be the case if we choose the λ_n so that $|\lambda_n| \leqq n^{-n}$. We are next going to use the fact that the field P is dense in the field R, and also that every imaginary field K is dense in the field C. We can, therefore, choose the λ_n, one after the other, within R or C, in such a way that (i) holds and that the number $b_n = P_n + \lambda_n$ is in P or in K, respectively. The construction of the desired function $g(z)$ is thereby completed.

5. **Divisibility in** $K\langle z\rangle$. On account of the previous theorem the divisibility structure of $K\langle z\rangle$ is found to be exceedingly simple. We have

THEOREM 4. *Let $f(z)$ be any element of $K\langle z\rangle$, and Z_f its zero-set. If K is real, there is in $K\langle z\rangle$, corresponding to every symmetric subset Z of Z_f, a divisor $g(z)$ of $f(z)$ with $Z_g = Z$. If K is imaginary, there is in $K\langle z\rangle$, corresponding to every subset Z of Z_f, a divisor $g(z)$ of $f(z)$ with $Z_g = Z$.*

This follows at once from Theorem 3. As an immediate consequence we have

THEOREM 5. *Let $f(z)$ be an element of $K\langle z\rangle$ that is neither zero nor a unit. If K is imaginary, $f(z)$ is irreducible if and only if $f(z)$ has exactly one root. If K is real, $f(z)$ is irreducible if and only if $f(z)$ has either exactly one real root, or exactly two roots which are conjugate imaginary.*

The following fact may be mentioned as a

COROLLARY. *Every function $f(z)$ of $K\langle z\rangle$ whose zero-set is infinite has infinitely many non-associated divisors in $K\langle z\rangle$.*

Thus, the divisor chain theorem does not hold in any $K\langle z\rangle$. To give a concrete example, of the functions

$$\sin z, \sin \tfrac{1}{2}z, \sin \tfrac{1}{4}z, \sin \tfrac{1}{8}z, \cdots \tag{7}$$

each is a proper divisor of the preceding, and all are contained in $P\langle z\rangle$ and hence in every $K\langle z\rangle$.

6. **The fundamental theorem in** $K\langle z\rangle$. Although the irreducible factors of an element $f(z)$ of $K\langle z\rangle$ are uniquely determined (apart from unit factors), it is not yet obvious that $f(z)$ can be represented as a product of them, for they are, in

general, infinite in number, and the question of convergence has to be considered. However, it turns out that we do have unique factorization in $K\langle z\rangle$:

THEOREM 6. *Every function $f(z)$ of $K\langle z\rangle$, which is neither zero nor a unit, can be represented as a (finite or infinite) product of irreducible functions of $K\langle z\rangle$. This representation is unique apart from the order of the factors and the occurrence of unit factors.*

Proof. Consider first the case where K is imaginary. Let

$$f(z) = C\cdot z^m\cdot e^{\psi(z)}\cdot \prod_k f_k(z) \tag{8}$$

be the Weierstrass representation of $f(z)$, where the product has at least one factor and may have infinitely many, and where each $f_k(z)$ has exactly one root. C may be chosen so that $\psi(0) = 0$; C is then the first non-vanishing coefficient of $f(z)$, and hence is in K.

By Theorem 3, every $f_k(z)$ can be multiplied by a unit $e^{\varphi_k(z)}$ so that $f_k(z)\cdot e^{\varphi_k(z)}$ has coefficients in K; here, if

$$\varphi_k(z) = \sum_{n=1}^{\infty} \lambda_{kn} z^n, \tag{9}$$

the $|\lambda_{kn}|$ must be $\leqq n^{-n}$ and may be chosen arbitrarily small; in particular, we may demand that

$$|\lambda_{kn}| \leqq \frac{1}{2^k\cdot n^n}. \tag{10}$$

This makes $\sum \varphi_k(z)$ convergent:

$$\varphi(z) = \sum_k \varphi_k(z) = \sum_k \sum_n \lambda_{kn} z^n = \sum_n z^n\cdot \sum_k \lambda_{kn} \tag{11}$$

with

$$\left|\sum_k \lambda_{kn}\right| \leqq \sum_k |\lambda_{kn}| \leqq \sum_k \frac{1}{2^k\cdot n^n} \leqq \frac{1}{n^n}, \tag{12}$$

so that $\varphi(z)$ is an integral function. If we set

$$g_k(z) = f_k(z)\cdot e^{\varphi_k(z)}, \tag{13}$$

we have

$$f(z) = C\cdot z^m\cdot e^{\psi(z)-\varphi(z)}\cdot \prod_k g_k(z). \tag{14}$$

If the product on the right is finite, it belongs to $K\langle z\rangle$, since each of its factors does; hence $e^{\psi-\varphi}$ belongs to $K\langle z\rangle$, and (14) is the desired representation of $f(z)$. This may not be so, if the product is infinite. In this case, let

$$\psi(z) - \varphi(z) = \sum_{n=1}^{\infty} \gamma_n z^n. \tag{15}$$

For each n, a sequence of numbers $\gamma_{1n}, \gamma_{2n}, \gamma_{3n}, \cdots$ may be chosen with the following three properties:

$$\gamma_{kn} \text{ in } K, \qquad \sum_k \gamma_{kn} = \gamma_n, \qquad |\gamma_{kn}| \leqq n^{-n}. \tag{16}$$

This is obviously possible since K is dense in C. Let

$$\psi_k(z) = \sum_{n=1}^{\infty} \gamma_{kn} z^n. \tag{17}$$

On account of (16), $\psi_k(z)$ is an element of $K\langle z\rangle$, and

$$\sum_k \psi_k(z) = \sum_n \gamma_n z^n = \psi(z) - \varphi(z). \tag{18}$$

Hence

$$\begin{aligned} f(z) &= C\cdot z^m \cdot \prod_k e^{\psi_k(z)} \cdot \prod_k g_k(z) \\ &= C\cdot z^m \cdot \prod_k [e^{\psi_k(z)} \cdot g_k(z)], \end{aligned} \tag{19}$$

where, by Theorem 1, every factor in the product is in $K\langle z\rangle$. Thus, we have also in this case arrived at the desired representation of $f(z)$. That the latter is unique apart from unit factors is obvious.

We have yet to deal with the case where K is real. The set Z_f is then symmetric, and $f(z)$ can be written in the form (8), where now, however, the $f_k(z)$ have a different meaning; they either have exactly one real root, or they have a couple of conjugate imaginary roots, so that (8) arises from the original Weierstrass product representation by a suitable grouping together of factors. In (8), $\psi(z)$ as well as every $f_k(z)$ is then in $R\langle z\rangle$, and the rest of the proof proceeds as before, except that the λ_{kn} must (and can) be chosen in R; in particular, (16) can be satisfied since K lies dense in R.

7. The Principal Ideal Theorem in $K\langle z\rangle$. The following theorem is immediately deducible from Theorem 3.

THEOREM 7. *Any finite or infinite set of functions of $K\langle z\rangle$ has a greatest common divisor in $K\langle z\rangle$.*

Proof. Let Z be the intersection of the zero-sets Z_f of the functions $f(z)$ under consideration. If K is real, every Z_f, and hence Z, is symmetric. Therefore, by Theorem 3, no matter whether K be real or imaginary, there will be a function $d(z)$ in $K\langle z\rangle$ for which $Z_d = Z$; this function is obviously the greatest common divisor of the functions $f(z)$. It is uniquely determined except for an arbitrary unit factor in $K\langle z\rangle$.

If d is the greatest common divisor of the functions $f, g, \cdots$, we shall write

$$d = \text{g.c.d.}\ (f, g, \cdots). \tag{20}$$

Theorem 7 is necessary, but not sufficient, for the validity of the Principal Ideal Theorem. In fact, the answer to the question whether there are none but principal ideals in the domain $K\langle z\rangle$ depends on the precise definition of "ideal" that is employed.

According to the classical definition, any subset of $K\langle z\rangle$ has to be called an *ideal* that has the following properties:

(a) if f and g are elements of the set, so is $f - g$;

(b) if f is an element of the set and h any element of $K\langle z\rangle$, then $h \cdot f$ is an element of the set.

Some ideals A have a *finite basis* $f_1, f_2, \cdots, f_n$, written

$$A = (f_1, f_2, \cdots, f_n), \tag{21}$$

in the sense that every element f of A can be written as a multiple sum of the f_i:

$$f = h_1 f_1 + h_2 f_2 + \cdots + h_n f_n, \tag{22}$$

where the h_i are in $K\langle z\rangle$. Among these ideals with a finite basis we have as a special case the *principal ideals* (f), whose basis consists of one element f.

In contradistinction to the classical definition, Prüfer[1] introduced a narrower ideal concept which comprises only the ideals (in the classical sense) with a finite basis. If an integral domain contains only ideals of this type, then the *Basis Theorem* is said to be valid in that domain. If, moreover, every ideal is a principal ideal, then the *Principal Ideal Theorem* is said to hold.

The answer to our question about the validity of the Principal Ideal Theorem in $K\langle z\rangle$ can now be formulated in the following two theorems.

THEOREM 8. *$P\langle z\rangle$, and hence every $K\langle z\rangle$, contains ideals (in the classical sense) which are not principal ideals.*

Proof. It had already been remarked in §5 above that in $K\langle z\rangle$ the Divisor Chain Theorem is not valid. The latter being equivalent to the Basis Theorem,[2] it follows at once that there are ideals in $K\langle z\rangle$ which do not have a finite basis, and hence are certainly not principal ideals. But without using this result, we can prove the theorem directly, by constructing an explicit counter-example.

We consider the ideal S generated by the chain of functions (7) (§5, above):

$$S = (\sin z, \sin \tfrac{1}{2}z, \sin \tfrac{1}{4}z, \cdots). \tag{23}$$

It is obvious that we have

$$z = \text{g.c.d. } (\sin z, \sin \tfrac{1}{2}z, \sin \tfrac{1}{4}z, \cdots) \tag{24}$$

and hence (z) is a divisor of S. However, S is not identical with the principal ideal (z). For z is not an element of S, since, if it were, it would have to be representable as a multiple sum of a finite number of the functions (7), say

$$z = h_1 \cdot \sin z + h_2 \cdot \sin \tfrac{1}{2}z + \cdots + h_n \cdot \sin 2^{-n+1}z. \tag{25}$$

[1] H. Prüfer, *Untersuchungen über Teilbarkeitseigenschaften in Körpern*, Journal für die reine und angewandte Mathematik, vol. 168(1932), pp. 1–36.

[2] Cf. B. L. van der Waerden, *Moderne Algebra*, II, §80.

But this is impossible, since $\sin 2^{-n+1}z$ is a divisor of the right side, and not of the left. On the other hand, S cannot be of the form $(z \cdot g(z))$, since, because of (24), no proper multiple of (z) divides S. S is, therefore, not a principal ideal. (As a result of the next theorem, it will appear that S does not even have a finite basis.)

THEOREM 9. *Every ideal in $K\langle z\rangle$ that possesses a finite basis is a principal ideal.*

If, therefore, with Prüfer, we should take only the ideals with finite bases into consideration, we should have to call $K\langle z\rangle$ a principal ideal ring. If we use the classical terminology, the content of Theorems 8 and 9 may also be expressed as follows. The ideals of $K\langle z\rangle$ fall into two non-empty classes; the ideals in one of these classes have the property that they possess only infinite bases; the ideals in the other class are all principal ideals; that is, they possess bases consisting of one element.

Proof of Theorem 9. The proof will amount to showing that the greatest common divisor of n functions $f_1, f_2, \cdots, f_n$ of $K\langle z\rangle$ can be represented as a multiple sum of these functions with coefficients in $K\langle z\rangle$. Let

$$\text{g.c.d. } (f_1, f_2, \cdots, f_n) = d. \tag{26}$$

Then, if there are $h_1, h_2, \cdots, h_n$ in $K\langle z\rangle$ such that

$$d = h_1 f_1 + h_2 f_2 + \cdots + h_n f_n, \tag{27}$$

it is obvious that the ideal $(f_1, f_2, \cdots, f_n)$ is a principal ideal:

$$(f_1, f_2, \cdots, f_n) = (d). \tag{28}$$

For, every f_i is contained in (d), and, on account of (27), d is contained in $(f_1, f_2, \cdots, f_n)$.

It will suffice to show (27) for $n = 2$. The general case then follows by induction: (i) for $n = 1$ we have $d = f_1$; (ii) if $d_1 = \text{g.c.d. } (f_1, f_2, \cdots, f_{n-1}) = j_1 f_1 + j_2 f_2 + \cdots + j_{n-1} f_{n-1}$, and $d = \text{g.c.d. } (d_1, f_n) = k_1 d_1 + k_2 f_n$, then $d = \text{g.c.d. } (f_1, f_2, \cdots, f_n)' = k_1 j_1 f_1 + \cdots + k_1 j_{n-1} f_{n-1} + k_2 f_n$, which is of the form (27).

Furthermore, if g.c.d. $(f_1, f_2) = d$, we have the equivalent relation $\left(\frac{f_1}{d}, \frac{f_2}{d}\right) = 1$, so that we may restrict ourselves to a pair of relatively prime functions of $K\langle z\rangle$:

$$\text{g.c.d. } (f_1, f_2) = 1. \tag{29}$$

We have to show that

$$(f_1, f_2) = (1) = K\langle z\rangle \tag{30}$$

or, what is the same, that there are two functions h_1 and h_2 in $K\langle z\rangle$ for which

$$1 = h_1 \cdot f_1 + h_2 \cdot f_2. \tag{31}$$

We first consider the case where K is an imaginary field. Since f_1 and f_2 are relatively prime, they do not both vanish at 0, and we may assume

$$(32)\qquad f_2(0) \neq 0.$$

Let $f_2(z)$ have the zeros $\beta_1, \beta_2, \cdots$ with multiplicities $r_1, r_2, \cdots$. If β stands for any one β_i, r being its multiplicity, we have an expansion for $f_1(z)\cdot f_2(z)$ at $z = \beta$ of the following form:

$$(33)\qquad f_1(z)\cdot f_2(z) = w_r(z-\beta)^r + w_{r+1}(z-\beta)^{r+1} + \cdots$$

with $w_r \neq 0$ because of (29). Further, let $M(z)$ be a meromorphic function whose only poles are poles of order r at the points β, with principal parts $m_r(z-\beta)^{-r} + m_{r-1}(z-\beta)^{-r+1} + \cdots + m_1(z-\beta)^{-1}$, the m_i being quantities to be determined presently. According to Mittag-Leffler, $M(z)$ is uniquely determined apart from an additive integral function. Thus, if $M_0(z)$ is any particular function of this kind, we have

$$(34)\qquad M(z) = M_0(z) + k(z),$$

where $k(z)$ is any integral function. Since $M_0(z)$ is regular at $z = 0$ on account of (32), it can be expanded into a power series in z. We now choose the coefficients of $k(z)$ so that $k(z)$ is an integral function and $M(z)$ has coefficients in K. This is clearly possible since K is assumed to be imaginary, and hence is dense in the field of all complex numbers.

We are still at liberty to choose the coefficients m_i of the principal parts of $M(z)$. We determine them, for each principal part, from the following equations:

$$(35)\qquad \begin{aligned} w_r m_r &= 1, \\ w_r m_{r-1} + w_{r+1} m_r &= 0, \\ \cdots\ \cdots\ \cdots\ \cdots\ &\cdots \\ w_r m_1 + w_{r+1} m_2 + \cdots + w_{2r-1} m_r &= 0. \end{aligned}$$

This system is solvable since its determinant is $-w_r^r \neq 0$.

The function $h_1(z)$ in (31) can now be determined as follows:

$$(36)\qquad h_1(z) = M(z)\cdot f_2(z).$$

$h_1(z)$ is clearly an integral function and hence lies in $K\langle z\rangle$, since both M and f_2 have coefficients in K. Moreover, because of (33) and (35), we have

$$(37)\quad h_1(z)f_1(z) = M(z)f_1(z)f_2(z) = 1 + c_r(z-\beta)^r + c_{r+1}(z-\beta)^{r+1} + \cdots.$$

Hence, for every β, $(z-\beta)^r$ is a divisor of $1 - h_1f_1$; in other words, f_2 is a divisor of $1 - h_1f_1$:

$$(38)\qquad 1 - h_1(z)f_1(z) = h_2(z)f_2(z)$$

and this is the same as (31). Here, the function $h_2(z)$ has automatically coefficients in K.

This finishes the proof, as far as an imaginary field K is concerned. If K is real, the set Z_{f_2} is symmetric, so that to every imaginary β we have a zero $\bar{\beta}$ of the same multiplicity r. For a real β, the w_i and hence the m_i will also be real. For a pair of conjugate imaginary roots β and $\bar{\beta}$, on the other hand, the corresponding w_i, and hence also the m_i, will be conjugate imaginary. It can be seen immediately from the customary proof of the Mittag-Leffler Theorem that under these conditions (symmetric set of poles, conjugate imaginary poles having "conjugate" principal parts) the function $M_0(z)$ can be made to have real coefficients. Whereupon $k(z)$ can again be chosen so that $M(z) = M_0(z) + k(z)$ has coefficients in K. In every other respect, the above proof remains literally the same.

8. **Integral functions of finite order.** The preceding considerations show that the divisibility structure of the domains $K\langle z\rangle$ is extremely simple, in fact, too simple to be conducive to the establishment of possible algebraic relations between the coefficients and the zeros of an integral function. In terms of the zeros we can distinguish whether the coefficients are real or imaginary, but can make no other inference about the algebraic nature of the coefficients. For R and C are the only fields which are "analytically closed" in the sense of containing all their accumulation points, and every K, according as it is real or imaginary, is dense either in R or in C; the latter property being, as has been seen, the only one that matters.

As a consequence, we have already in $P\langle z\rangle$ and in $P(i)\langle z\rangle$ decomposability into functions with at most two and with exactly one zero each, a situation which, as far as *polynomial* domains are concerned, arises only in the case of $R[z]$ and $C[z]$. This is all the more unsatisfactory, as the polynomial domain $K[z]$ is a subset of the domain $K\langle z\rangle$, so that a polynomial which is irreducible in $K[z]$ will in general become reducible if considered as an element of the wider domain $K\langle z\rangle$. For example, $z^2 - 2$ is irreducible in $P[z]$; but there must be a unit function $e^{\varphi(z)}$ providing the decomposition

$$z^2 - 2 = (z - \sqrt{2})e^{\varphi(z)} \cdot (z + \sqrt{2})e^{-\varphi(z)} \tag{39}$$

in $P\langle z\rangle$.

In view of this situation it is natural to try to replace $K\langle z\rangle$ by a more restricted domain of integral functions in which at least ordinarily irreducible polynomials remain irreducible. In the case of $P\langle z\rangle$, this restriction has to be of such a kind that, for example, the factor $(z - \sqrt{2})e^{\varphi(z)}$ in (39) belongs to the discarded functions of $P\langle z\rangle$. Now we shall see that a characteristic property of this factor is its being of an infinite order, that is to say, the function $(z - \sqrt{2})e^{\varphi(z)}$ can have rational coefficients only if $\varphi(z)$ is not a polynomial but a transcendental integral function. Thus, the factorization (39) can be eliminated by confining the investigation to the domains of integral functions of finite order with coefficients in a field K. In §2 we had already introduced the notation $K^*\langle z\rangle$ for these domains. The following theorems will show that

the transition from $K\langle z\rangle$ to $K^*\langle z\rangle$ really yields the desired result, namely, the irreducibility of ordinarily irreducible polynomials. On the other hand, this ad hoc restriction can hardly be said to be too artificial since, on account of their greater simplicity and the inclusion of the most frequently occurring integral functions among them, the integral functions of finite order have often in the literature been considered apart and made the subject of a separate investigation.

For the proof of the next theorem we shall require the following lemma which is in itself of some interest.

LEMMA. *Let $v_1, v_2, \cdots, v_n$ be distinct complex numbers, all different from zero; and let*

$$s_i = v_1^i + v_2^i + \cdots + v_n^i \qquad (i = 0, 1, 2, \cdots). \tag{40}$$

Then, if $s_{k+1}, s_{k+2}, s_{k+3}, \cdots$ are contained in the field K, so are $s_0, s_1, s_2, \cdots, s_k$.

Proof. Let c_i be the corresponding elementary symmetric functions of the v_i. If $m = \max(k + 1, n)$, we have for $i \geqq m$ the Newton relations

$$s_i c_n + s_{i+1}c_{n-1} + \cdots + s_{i+n-1}c_1 = -s_{i+n}. \tag{41}$$

These equations, for $i = m, m + 1, \cdots, m + n - 1$, form a system of n linear equations in $c_1, c_2, \cdots, c_n$, which can be written in matrix form, thus:

$$S_iC = T_i, \tag{42}$$

where

$$S_i = \begin{pmatrix} s_i & s_{i+1} & \cdots & s_{i+n-1} \\ s_{i+1} & s_{i+2} & \cdots & s_{i+n} \\ \cdots & \cdots & \cdots & \cdots \\ s_{i+n-1} & s_{i+n} & \cdots & s_{i+2n-2} \end{pmatrix}; \quad C = \begin{pmatrix} c_n \\ c_{n-1} \\ \cdots \\ c_1 \end{pmatrix}; \quad T_i = -\begin{pmatrix} s_{i+n} \\ s_{i+n+1} \\ \cdots \\ s_{i+2n-1} \end{pmatrix}. \tag{43}$$

Now let

$$V = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ v_1 & v_2 & \cdots & v_n \\ \cdots & \cdots & \cdots & \cdots \\ v_1^{n-1} & v_2^{n-1} & \cdots & v_n^{n-1} \end{pmatrix}; \qquad X = \begin{pmatrix} v_1 & 0 & \cdots & 0 \\ 0 & v_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & v_n \end{pmatrix}. \tag{44}$$

Then we have $VV' = S_0$, and more generally

$$VX^iV' = S_i. \tag{45}$$

$| V |^2 = d$ is the discriminant of the distinct quantities $v_1, v_2, \cdots, v_n$, and hence $\neq 0$; also $| X | = v_1v_2 \cdots v_n = c_n \neq 0$, so that

$$| S_i | = d \cdot c_n^i \neq 0. \tag{46}$$

We may, therefore, write

$$C = S_i^{-1}T_i. \tag{47}$$

But this means that the c_i, and hence also all the s_i, are expressible as rational functions of the elements of S_i and T_i; these elements being just the s_{k+1}, s_{k+2}, $\cdots$ which by assumption are in K. It follows that all s_i are elements of K.

We are now in a position to prove the following.

THEOREM 10. *Let K_2 be a field over K_1, and let $F(z)$ be a polynomial of $K_2[z]$ but not of $K_1[z]$, and let $F(0) = 1$. Then, for every function $f(z)$ of $K_1\langle z\rangle$ with $Z_f = Z_F$, we have* $\operatorname{ord} f = \infty$.

We know by Theorem 3 that there exists a function $f(z)$ in $K_1\langle z\rangle$ for which $Z_f = Z_F$, unless K_1, but not K_2, is real. We have to show that $f(z)$ cannot be of a finite order.

Proof. We may assume $f(0) = 1$. Since $Z_f = Z_F$, we must have

$$F(z)\cdot e^{\varphi(z)} = f(z). \tag{48}$$

Suppose now that $f(z)$ is of finite order; then $\varphi(z)$ must be a polynomial in $K_2[z]$:

$$\varphi(z) = \lambda_1 z + \lambda_2 z^2 + \cdots + \lambda_k z^k \tag{49}$$

with λ_i in K_2. Let $F(z)$ have the zeros $-\zeta_1, -\zeta_2, \cdots, -\zeta_n$, so that

$$F(z) = \left(1 + \frac{z}{\zeta_1}\right)\left(1 + \frac{z}{\zeta_2}\right)\cdots\left(1 + \frac{z}{\zeta_n}\right). \tag{50}$$

If, now, we set

$$s_i = \left(\frac{1}{\zeta_1}\right)^i + \left(\frac{1}{\zeta_2}\right)^i + \cdots + \left(\frac{1}{\zeta_n}\right)^i, \tag{51}$$

we obtain

$$\begin{aligned}\log F(z) &= \sum_{i=1}^{n} \log\left(1 + \frac{z}{\zeta_i}\right) = \sum_{i=1}^{n}\left[\frac{z}{\zeta_i} - \frac{1}{2}\left(\frac{z}{\zeta_i}\right)^2 + \frac{1}{3}\left(\frac{z}{\zeta_i}\right)^3 - \cdots\right] \\ &= s_1 z - \tfrac{1}{2}s_2 z^2 + \tfrac{1}{3}s_3 z^3 - + \cdots.\end{aligned} \tag{52}$$

On the other hand, $f(z)$ has the form

$$f(z) = 1 + a_1 z + a_2 z^2 + \cdots \tag{53}$$

with a_i in K_1. Hence

$$\log f(z) = b_1 z + b_2 z^2 + \cdots, \tag{54}$$

where the b_i are polynomials in the a_i with rational coefficients, and hence themselves in K_1. We now have, by taking logarithms in (48),

$$\begin{aligned}&\lambda_1 z + \lambda_2 z^2 + \cdots + \lambda_k z^k + s_1 z - \tfrac{1}{2}s_2 z^2 + \tfrac{1}{3}s_3 z^3 - \cdots \\ &\qquad = b_1 z + b_2 z^2 + b_3 z^3 + \cdots.\end{aligned} \tag{55}$$

Comparing coefficients we find that, for $i > k$, the s_i are in K_1; hence, by the above lemma, all s_i are in K_1, and thus also the λ_i. But this would imply

that $F(z)$ has coefficients in K_1 and contradict our assumption. The supposition that $f(z)$ is of finite order is thereby refuted.

As an immediate consequence of Theorem 10 we have the theorem on the irreducibility of polynomials.

THEOREM 11. *If $F(z)$ is an irreducible polynomial of $K[z]$, then it is also irreducible in $K^*\langle z\rangle$.*

Proof. Suppose $F(z)$ factors in $K^*\langle z\rangle$:

$$F(z) = f_1(z)\cdot f_2(z), \tag{56}$$

where neither $f_i(z)$ is a unit. Let

$$F(z) = F_1(z)\cdot F_2(z) \tag{57}$$

be the corresponding factorization in $K_2[z]$, where K_2 is the root field of $F(z)$ over $K_1 = K$. Then

$$f_i(z) = F_i(z)\cdot e^{\varphi_i(z)} \qquad (i = 1, 2), \tag{58}$$

where the $\varphi_i(z)$ are polynomials, since the $f_i(z)$ are of finite order. From this, by Theorem 11, it follows that the $F_i(z)$ themselves have coefficients in K, so that $F(z)$ is reducible in $K[z]$.

A further investigation into the properties of the domains $K^*\langle z\rangle$ will be made in another article. In particular, it will be necessary to deal with such problems as have above been answered with reference to the domains $K\langle z\rangle$. Does, for example, the fundamental theorem apply in $K^*\langle z\rangle$? Is the greatest common divisor of two elements representable as a multiple sum of those elements? Are there any simple criteria of irreducibility?

UNIVERSITY OF ILLINOIS.

INFINITE POWERS OF MATRICES AND CHARACTERISTIC ROOTS

By Rufus Oldenburger

1. **Introduction.** Frazer, Duncan and Collar [2][1] have studied special cases of infinite powers of complex matrices. Necessary and sufficient conditions for the existence of such powers and the vanishing of such powers are given in the present paper. By the use of these powers one can solve one of the outstanding problems in applied mechanics [1, 6]. The results for vanishing infinite powers yield a simple proof of a theorem of Frobenius [3] to the effect that when the complex matrices are partially ordered in a certain way, the maximum absolute value of the characteristic roots of a complex matrix A remains the same or increases when A is increased. This result is useful in computing upper limits to the characteristic roots of a matrix, especially for real matrices with non-negative elements. Such matrices have been studied by various writers [4, 7].

2. **Definitions.** Let A be a square matrix with elements in the complex field. Let the elements in the i-th row and j-th column of A^n be denoted by $a_{ij}^{[n]}$. If for each pair of values of i and j the elements $a_{ij}^{[1]}, a_{ij}^{[2]}, a_{ij}^{[3]}, \cdots$ converge to a limit b_{ij} in the usual sense, we say that A^∞ *exists* and is the matrix (b_{ij}). Otherwise, we say that A^∞ does not exist. Let $A = (a_{ij})$ and $B = (b_{ij})$ have respectively non-negative and complex elements which satisfy $a_{ij} \geqq | b_{ij} |$. We say that A *contains* B, and write $A \supset B$. If $A \supset 0$, we term A *non-negative.* The smallest circle with center at the origin in the Argand diagram containing the characteristic roots of A is termed the *characteristic circle* of A. Let $[A]$ be the matrix whose elements are the absolute values of the corresponding elements of A. We term $[A]$ the *absolute matrix* of A.

3. **Existence of the infinite power of a matrix.** By means of the Jordan normal form[2] we shall prove the following

Theorem 1. *The infinite power of a complex matrix A exists if and only if the characteristic roots of A corresponding to elementary divisors (taken with respect to the complex field) of degree greater than* 1 *are in absolute value less than* 1, *and the remaining characteristic roots are equal to* 1 *or in absolute value less than* 1.

Received September 1, 1939; in revised form, November 29, 1939; presented to the American Mathematical Society, September 8, 1939.

[1] Numbers in brackets refer to the bibliography at the end of the paper.

[2] L. E. Dickson, *Modern Algebraic Theories*, p. 106.

Since $X^{-1}A^nX = (X^{-1}AX)^n$ for non-singular X, the power A^∞ exists if and only if $(X^{-1}AX)^\infty$ exists. Choose X so that $X^{-1}AX$ is in the Jordan normal form C, where

$$C = \begin{Vmatrix} C_1 & 0 & \cdots & 0 \\ 0 & C_2 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & C_m \end{Vmatrix},$$

and further the minor C_i given by

$$C_i = \begin{Vmatrix} \lambda_i & 0 & 0 & \cdots & \cdots & \cdots & 0 \\ 1 & \lambda_i & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & 1 & \lambda_i & \cdots & \cdots & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & \cdots & 0 & 1 & \lambda_i \end{Vmatrix}$$

corresponds to one elementary divisor of $A - \lambda I$ for each i. Here I denotes the identity matrix. Now C^∞ exists if and only if C_i^∞ exists for each i. Let C_i be of order ρ. Then for $n \geqq \rho - 1$

$$C_i^n = \begin{Vmatrix} \lambda_i^n & 0 & 0 & 0 & \cdots & 0 \\ a_1\lambda_i^{n-1} & \lambda_i^n & 0 & 0 & \cdots & 0 \\ a_2\lambda_i^{n-2} & a_1\lambda_i^{n-1} & \lambda_i^n & 0 & \cdots & 0 \\ a_3\lambda_i^{n-3} & a_2\lambda_i^{n-2} & a_1\lambda_i^{n-1} & \lambda_i^n & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{\rho-1}\lambda_i^{n-\rho+1} & \cdots & \cdots & \cdots & \cdots & \cdots \end{Vmatrix},$$

where $a_j = C(n, j)$ for each j. It is easy to see that

$$\lim_{n\to\infty} a_j\lambda_i^{n-j} \qquad (j = 1, 2, \cdots, \rho - 1)$$

exists if and only if $|\lambda_i| < 1$. If $|\lambda_i| < 1$, then $C_i^\infty = 0$. If $|\lambda_i| = 1$, the limit of λ_i^n as $n \to \infty$ exists if and only if $\lambda_i = 1$.

Since the field of real numbers is closed with respect to the limit process, A^∞ is real if A is real and A^∞ exists. If A^∞ exists, the normal form of A^∞ is clearly

$$\begin{Vmatrix} I & 0 \\ 0 & 0 \end{Vmatrix},$$

where I is an identity matrix. It follows that the characteristic roots of A^∞ are zero or one.

The above proof yields also

THEOREM 2. *The infinite power of a complex matrix is zero if and only if the characteristic circle of A lies within the unit circle.*

4. **Characteristic circles of matrices.** By means of Theorem 2 we shall exhibit a simple proof of the following theorem, first proved by Frobenius in another manner.

THEOREM 3. *Let $A = (a_{ij})$ and $B = (b_{ij})$ be matrices with elements in the complex field. If A contains B, the characteristic circle of A contains the characteristic circle of B.*

We let λ and μ denote respectively the characteristic roots of A and B with maximum absolute value. We assume first that $\lambda \neq 0$. For arbitrary ϵ we can choose a real positive number α_ϵ so that the maximum absolute value of the characteristic roots of $\alpha_\epsilon A$ is $(1 - \epsilon)$. This maximum absolute value is $\alpha_\epsilon \mid \lambda \mid$. Clearly $A \supset [B] \supset 0$, whence $\alpha_\epsilon A \supset [\alpha_\epsilon B] \supset 0$. If $A \supset [B]$ and $A' \supset [B']$, then $AA' \supset [BB']$. Thus $(\alpha_\epsilon A)^n \supset [(\alpha_\epsilon B)^n] \supset 0$. Now

$$\lim_{n\to\infty} (\alpha_\epsilon A)^n = 0 \tag{1}$$

implies that

$$\lim_{n\to\infty} (\alpha_\epsilon B)^n = 0. \tag{2}$$

By Theorem 2 formula (1) is valid, whence by (2) the characteristic roots of $\alpha_\epsilon B$ lie within the unit circle with center at the origin. The maximum absolute value of the characteristic roots of $\alpha_\epsilon B$ is $\alpha_\epsilon \mid \mu \mid$, whence $\alpha_\epsilon \mid \mu \mid < 1$. As $\epsilon \to 0$, we have $\alpha_\epsilon \to 1/\mid \lambda \mid$. Thus $\mid \mu \mid \leqq \mid \lambda \mid$.

If $\lambda = 0$, each characteristic root of A is zero, whence the characteristic equation of A is $\lambda^n = 0$. Thus $A^n = 0$. Since $A^n \supset [B^n]$, we have $B^n = 0$. From this it follows that $\mu = 0$.

Frobenius proved [3] that if the elements of a matrix A are all positive, the root λ is positive and simple. If A is non-negative but has some zero elements, we replace each zero element by ϵ to obtain from A a matrix A_ϵ. If $\epsilon > 0$ the result of Frobenius implies that A has a positive characteristic root λ_ϵ greater than the absolute value of any other. Since by Theorem 3 the inequality $\epsilon \geqq \epsilon'$ implies that $\lambda_\epsilon \geqq \lambda_{\epsilon'}$, we have

$$\lim_{\epsilon\to 0} \lambda_\epsilon \geqq 0.$$

Thus λ is *non-negative for each non-negative A.*

We remark that the only matrices compared with a complex matrix A are the absolute matrix $[A]$ of A and non-negative matrices obtained by increasing the elements of $[A]$.

In view of Theorem 2, Theorem 3 is of considerable computational value in determining the vanishing of the infinite power of a given complex matrix.

THEOREM 4. *Let B be a non-negative matrix obtained from a non-negative matrix A by bordering A by rows and columns. The characteristic circle of B is as great as the characteristic circle of A.*

We border A by a row and column of zeros to obtain the matrix A'. Since the characteristic roots of A' are the characteristic roots of A and zero, the theorem holds for A and $B = A'$. Any matrix B obtained from A by bordering A by a row and column of non-negative elements satisfies the relation

$$B \supset A'.$$

By Theorem 3, Theorem 4 is valid for this choice of A and B. The proof of the theorem for any B now follows by induction.

5. **A theorem of Hirsh.** The following theorem, proved by Hirsh [5] in another manner, is a consequence of Theorem 3.

THEOREM 5. *Let M be the least upper bound of the elements of a matrix A of order n with elements in the complex field. The characteristic roots of A are less than or equal to nM in absolute value.*

By Theorem 3 the maximum absolute value of the characteristic roots of A is not greater than this maximum for the matrix B of order n with all elements equal to M. Let C denote the matrix of order n all of whose elements are equal to one. The characteristic roots of B are evidently the products of the characteristic roots of C by M. The maximum positive characteristic root of C is n. Hence the maximum absolute value of the characteristic roots of B is nM.

6. **Algebraic equations.** Except for sign we can write the polynomial $P(\lambda) = \lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + \cdots + a_n$ as the characteristic determinant of the matrix

$$A = \left\| \begin{matrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & 1 \\ (-a_n) & \cdots & \cdots & \cdots & \cdots & (-a_2) & (-a_1) \end{matrix} \right\|.$$

We use the expression *circle of roots for* $P(\lambda) = 0$ for the characteristic circle of A. Theorem 3 implies the known

THEOREM 6. *If* $|a_i| \leqq b_i$ *for every* i, *the circle of roots for*

$$\lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n = 0$$

is contained in the circle of roots for

$$\lambda^n - b_1\lambda^{n-1} - \cdots - b_{n-1}\lambda - b_n = 0.$$

7. On Theorem 3. Professor Weyl has communicated to the author a second proof of Theorem 3 which we sketch here. For power series we use the notation

$$\alpha_1 + \alpha_2 t + \alpha_3 t^2 + \cdots \subset \beta_1 + \beta_2 t + \beta_3 t^2 + \cdots$$

to mean that $|\alpha_i| \leqq \beta_i$ for each i. Now $\phi_A(t) = \text{trace}\, A(I - tA)^{-1}$ is a power series in t. Clearly, $A \supset B$ implies $\phi_A(t) \supset \phi_B(t)$. Thus the circle of convergence of $\phi_B(t)$ contains the circle for $\phi_A(t)$. If ρ_A is the maximum absolute value of the characteristic roots of A, the radius of convergence of $\phi_A(t)$ is ρ_A^{-1}. It follows that $\rho_A \geqq \rho_B$. As Professor Weyl remarks, these results hold for a large class of operators in Hilbert space, and in particular for Fredholm's integral equations.

Bibliography

1. H. Cross, *Analysis of continuous frames*, Transactions of the American Society of Civil Engineers, vol. 96(1932), pp. 1–10.
2. R. A. Frazer, W. J. Duncan, and A. R. Collar, *Elementary Matrices*, Cambridge University Press, 1938.
3. S. B. Frobenius, *Über Matrizen aus positiven Elementen*, Preuss. Akad. Wiss., Sitzungsberichte, 1909, pp. 514–518.
4. S. B. Frobenius, *Über Matrizen aus positiven Elementen*, II, Preuss. Akad. Wiss., Sitzungsberichte, 1908, pp. 471–476.
5. A. Hirsh, *Sur les racines d'une équation fondamentale*, Acta Mathematica, vol. 25(1902), pp. 367–370.
6. R. Oldenburger, *Convergence of the Hardy Cross balancing process*, Journal of Applied Mechanics, 1940.
7. M. Potron, *Sur les matrices non-négatives*, Comptes Rendus, vol. 204(1937), pp. 844–846.

Armour Institute of Technology.

THE PARTIAL DIFFERENTIAL EQUATION $\frac{\partial z}{\partial x} + f(x, y)\frac{\partial z}{\partial y} = 0$

By Louis D. Rodabaugh

Theorem 1. *Let g be a bounded, open, simply connected plane region. Let (x, y) be the rectangular Cartesian coördinates of a general point P of the plane for a particular coördinate system. Let $f(x, y)$ be a function*[1] *such that*

(a) *$f(x, y)$ and $f_y(x, y)$ are defined and are continuous in g;*

(b) *$f(x, y)$ and $f_y(x, y)$ have definite finite continuous limits on the boundary of g. Then there exists a function $\psi(x, y)$ such that in g:*

(a) *$\psi(x, y)$ is defined and is of class C′ with respect to x and y;*

(b) *$\psi(x, y)$ satisfies*

$$\frac{\partial z}{\partial x} + f(x, y)\frac{\partial z}{\partial y} = 0; \tag{1}$$

(c) *$\psi(x, y)$ satisfies*

$$\psi_y(x, y) > 0. \tag{2}$$

Proof. A solution curve of

$$y' = f(x, y) \tag{3}$$

shall be called a "characteristic" of (1). A known theorem assures that through each point of g there passes exactly one characteristic of (1), that these characteristics approach arbitrarily close to the boundary of g in both directions of the x-axis, and that they depend continuously on the initial point.

Remarks made by Kamke[2] for the case where g is an open, simply connected region "lying entirely in" an open region G apply also to the present case, and these remarks constitute the proof of

Lemma 1. *There exists a set of open, simply connected regions $q_1, q_2, \cdots$ with the properties:*

(a) *each q_n [$\equiv q(t_n)$] is the set of points belonging to characteristics of* (1) *which lie in g and pass through an open finite vertical segment t_n lying in g;*

(b) *each h_n [$\equiv q_1 + \cdots + q_n$] is an open region;*

(c) *$q_1 + q_2 + \cdots = g$;*

(d) *the common points of t_n and h_{n-1} form exactly one open segment, and t_n projects out of h_{n-1} in exactly one direction.*

Received September 1, 1939.

[1] Only real functions of real variables are considered in this paper.

[2] [1], §2, pp. 605–609. The numbers in brackets refer to the bibliography

Denote by x_ν the abscissa of t_ν, and by $y = \varphi(x, \xi, \eta)$ the characteristic of (1) passing through the point (ξ, η) of g. Then (by a theorem of Bendixson) the function $\psi(\xi, \eta) \equiv \varphi(x_1, \xi, \eta)$ possesses for h_1 [$\equiv q(t_1)$] rather than for g (and with ξ, η instead of x, y) the properties described in the conclusion of Theorem 1.

Assume that a function $\psi(\xi, \eta)$ with these properties has already been constructed for h_{n-1}, and that

$$\psi(\xi, \eta) \equiv \omega_\nu[\varphi(x_\nu, \xi, \eta)] \qquad \text{in } q(t_\nu)\ (\nu < n),$$

where $\omega_\nu(y)$ is of class C′ on t_ν, $\omega_\nu'(y) > 0$ on t_ν, and $\omega_\nu'(y)$ has definite finite positive limits at the endpoints of t_ν.

The segment t_n projects out of h_{n-1} in just one direction, assume for definiteness the upper. Let V_n and W_n denote the ordinates of the endpoints of the

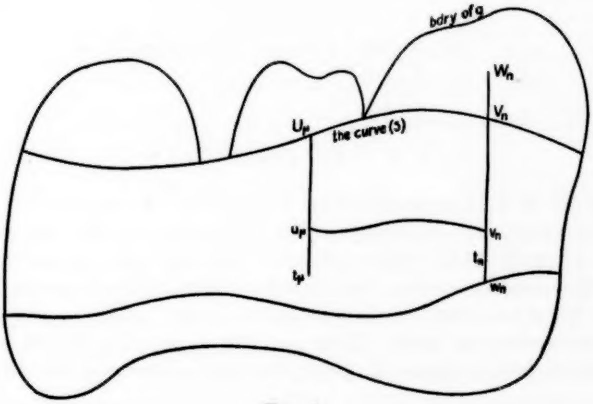

FIG. 1

half-open segment of t_n projecting outside h_{n-1}. Denote this interval simply by $\langle V_n, W_n)$, and proceed correspondingly in similar cases, as indicated in Figure 1. The lower segment of t_n, all the way to its lower endpoint w_n, belongs to h_{n-1}. This segment has exactly one open segment, if any points at all, in common with each of the regions $q_1, \cdots, q_{n-1}$. Hence there exists a smallest μ $(1 \leqq \mu < n)$ such that for some $v_n < V_n$ the open segment (v_n, V_n) of t_n belongs to q_μ and also to $h_\mu - h_{\mu-1}$. The characteristics of (1) through the open interval (v_n, V_n) therefore run through t_μ and cut out on this segment an open interval (u_μ, U_μ).[3]

The characteristics of (1),

$$y = \varphi(x, x_\mu, \eta), \tag{4}$$

[3] This paragraph is taken from [1], §3.3, p. 610.

through the interval (u_μ, U_μ) are of class C′, their slopes are bounded, and they satisfy (3). Hence, as η varies monotonically from u_μ to U_μ, the characteristics (4) converge uniformly to a curve

$$y = \varphi(x, x_\mu, U_\mu) \tag{5}$$

which is continuous and single-valued, and which is defined at least for $x_\mu \leqq x \leqq x_n$ (assume for simplicity that $x_\mu < x_n$). Every point of the curve (5) is in the closure of g. Since $f(x, y)$ has a definite finite continuous limit on the boundary of g, if the domain of definition of $f(x, y)$ is continuously extended to the closure of g, the curve (5) is itself a solution curve of (3), that is, a characteristic of (1).

By a theorem of Bendixson the function $\varphi(x_\mu, \xi, \eta)$ has $q(t_\mu)$ for its exact region of definition, is constant for all points of one and the same characteristic of (1), and is of class C′ with respect to ξ and η. Further,

$$\varphi_\eta(x_\mu, \xi, \eta) = \exp\left\{\int_\xi^x f_y[x, \varphi(x, \xi, \eta)]\, dx\right\} > 0, \tag{6}$$

and

$$\varphi_\xi(x_\mu, \xi, \eta) + f(\xi, \eta)\varphi_\eta(x_\mu, \xi, \eta) \equiv 0 \qquad \text{in } q(t_\mu). \tag{7}$$

Let the point (ξ, η) approach a point $P{:}\,(x, y)$ on the curve (5) from within h_{n-1}. The function $\varphi(x_\mu, \xi, \eta)$ approaches the limit U_μ. The function $f_\eta(\xi, \eta)$ approaches a definite finite continuous limit. By (6), $\varphi_\eta(x_\mu, \xi, \eta)$ approaches a definite finite continuous limit: this limit is positive, since the integrand in the exponent in (6) is bounded. It is clear from (7) that $\varphi_\xi(x_\mu, \xi, \eta)$ approaches a definite finite continuous limit. Thus $\varphi(x_\mu, \xi, \eta)$, $\varphi_\xi(x_\mu, \xi, \eta)$ and $\varphi_\eta(x_\mu, \xi, \eta)$ all have definite finite continuous limits, the last positive, on the curve (5).

Since

$$\psi(\xi, \eta) \equiv \omega_\mu[\varphi(x_\mu, \xi, \eta)] \qquad \text{in } q(t_\mu), \tag{8}$$

$\psi(\xi, \eta)$ has a definite finite continuous limit on the curve (5). As the point (ξ, η) approaches the point (x, y) on the curve (5) from within $q(t_\mu)$, the function $\psi_\eta(\xi, \eta)$ approaches

$$\{\lim_{y \to U_\mu} \omega_\mu'(y)\} \cdot \{\lim_{(\xi,\eta) \to (x,y)} \varphi_\eta(x_\mu, \xi, \eta)\}; \tag{9}$$

hence $\psi_\eta(\xi, \eta)$ has a definite finite continuous limit on the curve (5). This limit is positive, since each limit in (9) is positive. Similarly it is seen that $\psi_\xi(\xi, \eta)$ has a definite finite continuous limit on the curve (5).

There exists a function $\omega_n(y)$ such that

(a) $\omega_n(y)$ is defined and is of class C′ on $w_n < y < W_n$;

(b) $\omega_n'(y) > 0$ on $w_n < y < W_n$, and $\omega_n'(y)$ has definite finite positive limits at $y = w_n$ and $y = W_n$ (in particular, if M_n and M_n' are any finite real numbers

such that $M_n > \omega_n(V_n)$ and $M'_n > 0$, there exists a function $\omega_n(y)$ as described such that

$$\lim_{y \to W_n} \omega_n(y) = M_n \quad \text{and} \quad \lim_{y \to W_n} \omega'_n(y) = M'_n);$$

(c) $\omega_n(y) \equiv \psi(x_n, y)$ on $w_n < y < V_n$.

Consider the function $\psi_n(\xi, \eta)$ defined in $q(t_n)$ by

$$\psi_n(\xi, \eta) \equiv \omega_n[\varphi(x_n, \xi, \eta)]. \tag{10}$$

This function is of class C′ in $q(t_n)$, and also

(a) $\psi_{n\eta}(\xi, \eta) = \omega'_n[\varphi(x_n, \xi, \eta)]\cdot\varphi_\eta(x_n, \xi, \eta) > 0$ in $q(t_n)$;

(b) in that part of $q(t_n)$ which is contained in h_{n-1},

$$\psi_n(\xi, \eta) \equiv \omega_n[\varphi(x_n, \xi, \eta)] \equiv \psi[x_n, \varphi(x_n, \xi, \eta)] \equiv \psi(\xi, \eta); \tag{11}$$

(c) in $q(t_n)$,

$$\begin{aligned} &\psi_{n\xi}(\xi, \eta) + f(\xi, \eta)\psi_{n\eta}(\xi, \eta) \\ &\qquad = \omega'_n[\varphi(x_n, \xi, \eta)]\cdot\{\varphi_\xi(x_n, \xi, \eta) + f(\xi, \eta)\varphi_\eta(x_n, \xi, \eta)\} = 0, \end{aligned} \tag{12}$$

so that $\psi_n(\xi, \eta)$ is a solution of (1) in $q(t_n)$ (with ξ, η instead of x, y).

Thus the function $\psi(\xi, \eta)$ defined by

$$\psi(\xi, \eta) \equiv \omega_\nu[\varphi(x_\nu, \xi, \eta)] \qquad \text{in } q(t_\nu)\ (\nu = 1, \cdots, n) \tag{13}$$

possesses in h_n instead of g (and with ξ, η instead of x, y) the properties described in the conclusion of Theorem 1, and the induction is complete.

Remark 1. Only a denumerable number of steps are involved in the proof. It therefore follows from the parenthetical remark in property (b) of the function $\omega_n(y)$ that there exists a function $\psi(x, y)$ as described in Theorem 1 which is bounded in g. Further, if a (>0) and b are arbitrary real numbers and if $\psi(x, y)$ is a function as described in Theorem 1, then $\Psi(x, y)$ [$\equiv a\psi(x, y) + b$] is also such a function. It follows that there exists a function $\psi(x, y)$ as described in Theorem 1 which has the additional property

(d) $L < \psi(x, y) < U$ in g,

where L and U are any preassigned finite real numbers such that $L < U$.

Remark 2. Examples show that part (b) of the hypothesis of Theorem 1 is not necessary in order that there exist a function $\psi(x, y)$ as described, but from a theorem of Ważewski[4] it is clear that if part (b) is omitted the remaining conditions do not imply the existence of such a function.

Remark 3. From a theorem of Whitney[5] it is seen that if $f(x, y)$ is "of class C′ in a closed subset A of the xy-plane in terms of" certain functions, then there is a function $F(x, y)$ of class C′ in the whole xy-plane in the ordinary sense, such that $F(x, y) \equiv f(x, y)$ in A. Theorem 1 lies "between" Kamke's theorem[6]

[4] [2], p. 104.
[5] [3], Theorem 1, p. 65.
[6] [1], Satz I, p. 603.

and the existence theorem implied by Whitney's work, in this sense: The hypotheses on $f(x, y)$ and the region g in Theorem 1 are stronger than those on $f(x, y)$ and the region G in Kamke's theorem, but the solution of (1) obtained in Theorem 1 is defined all the way to the boundary of g. The hypotheses on $f(x, y)$ and the region g in Theorem 1 are weaker than those necessary in order to apply Whitney's theorem, but under his conditions a solution $\psi(x, y)$ of (1) in g exists for which there is a function $\Psi(x, y)$ of class C′ in the whole xy-plane such that $\Psi(x, y) \equiv \psi(x, y)$ in g.

THEOREM 2. *Let R be an open, doubly connected plane region lying between two simple closed curves K and k such that k is contained in the open region enclosed by K. Let (x, y) be the rectangular Cartesian coördinates of a general point P of the plane for a particular coördinate system. Let $f(x, y)$ be a function such that*

(a) *$f(x, y)$ is defined and is continuous in the closure of R and is of class* C′ *with respect to y in R;*

(b) *$f_y(x, y)$ has a definite finite continuous limit on the boundary of R;*

(c) *there is a leftmost point A of k through which there do not pass two distinct characteristics of* (1), *each having a limit point on K to the right of A; there is a rightmost point B of k through which there do not pass two distinct characteristics of* (1), *each having a limit point on K to the left of B.*

Then there exists a function $\psi(x, y)$ such that in R:

(a) *$\psi(x, y)$ is defined and is of class* C′ *with respect to x and y;*

(b) *$\psi(x, y)$ satisfies* (1);

(c) *$\psi(x, y)$ satisfies* (2).

Proof. Since $f(x, y)$ is bounded, a characteristic of (1) extends from A to the left. Denote by E the rightmost limit point of this characteristic which is on K and to the left of A. Denote by C_A the open segment EA of this characteristic. There exists a corresponding open segment C_B ($\equiv BF$) of a characteristic extending from B to the right.

Denote by s_A an open vertical segment which, including its endpoints, lies in R and which cuts C_A in a point Q. Denote by s_B a corresponding open vertical segment cutting C_B. Form two open vertical segments s_{AU} and s_{AL}, upper and lower respectively, by deleting Q from s_A. *There is at least one such segment so small that every characteristic cutting it has its left endpoint on K and its right endpoint on k.* Suppose that no matter how small s_{AU} is there is a characteristic which passes through s_{AU} and has its right endpoint on K. If m denotes the ordinate through A, there is a lowest point P_1 which lies above A and on K and m. Since characteristic-segments which lie in R depend continuously on the initial point, there is a characteristic through s_{AU} which has both endpoints on K and which passes through the open segment AP_1. Similarly, if P_2 denotes the midpoint of AP_1, there is one such characteristic which passes through the open segment AP_2. Similarly for AP_4, if P_4 denotes the midpoint of AP_2. And so on, indefinitely. Since the slopes of the characteristics are bounded, there can be determined in this manner a sequence of characteristics such that

(a) each characteristic passes through s_{AU};

(b) each characteristic has both endpoints on K;

(c) each characteristic extends at least a certain positive distance α to the right of A and has part of k below it;

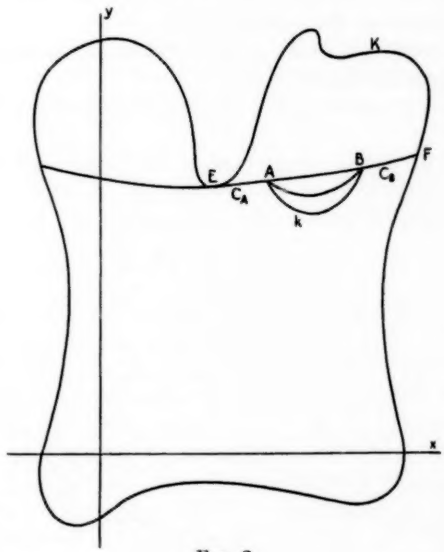

Fig. 2

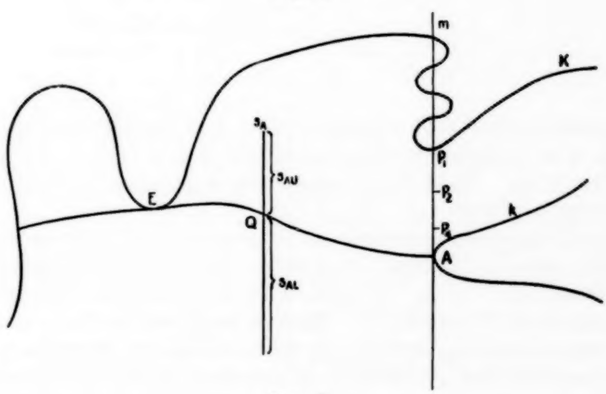

Fig. 3

(d) the sequence is uniformly convergent to a limiting curve C_U which contains C_A as a subsegment.

C_U is itself a characteristic, and C_U passes through A to the right on or above k and has a limit point on K to the right of A. If, no matter how small s_{AL}

is, there is a characteristic which passes through s_{AL} and has its right endpoint on K, then there is a characteristic C_L which passes through A to the right on or below k and has a limit point on K to the right of A. This contradicts part (c) of the hypothesis of the theorem.

Thus there exists a closed vertical segment t_A such that

(a) t_A is a subsegment of the segment s_A;

(b) through each point of t_A there passes a characteristic whose left endpoint is on K and whose right endpoint is on k.

There is also a segment t_B having the corresponding properties with reference to s_B.

Denote by l_1 and l_2 characteristics through the upper and lower endpoints, respectively, of t_A. Denote by G the leftmost limit point of l_1 on k, and by H the rightmost limit point of l_1 on K and to the left of G. Let U and V denote

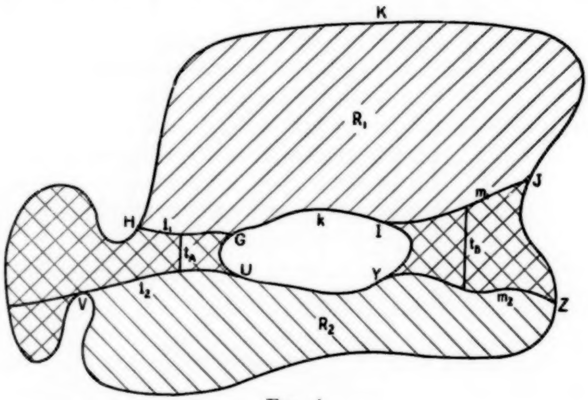

FIG. 4

the corresponding points with reference to l_2. Assume, without loss of generality, that t_A is so small and so placed on s_A that H and V are to the left of the ordinate through t_A. Make the corresponding denotations with reference to t_B, as indicated in Figure 4.

Denote by R_1 the open region enclosed by the curve which passes from U to Y in a positive sense[7] on k, from Y to Z on m_2, from Z to V in a positive sense on K, and from V to U on l_2. Denote by R_2 the open region enclosed by the corresponding curve $IYUGHVZJI$. Each of these regions is an open, simply connected region satisfying with $f(x, y)$ the hypothesis of Theorem 1. Hence there exist functions $\psi_1(x, y)$ and $\psi_2(x, y)$ possessing, in R_1 and R_2 respectively, the properties described in the conclusion of Theorem 1 and also satisfying

$$\psi_2(x, y) < M < N < \psi_1(x, y), \tag{14}$$

where M and N are any finite real numbers such that $M < N$.

[7] A point moves in a positive sense on the boundary of R if it moves so that the region R is always to its left.

Let $VUGH$ denote the open region enclosed by the curve which passes from V to U on l_2, from U to G in a positive sense on k, from G to H on l_1, and from H to V in a positive sense on K. Let $YZJI$ denote the corresponding open region with reference to t_B. Denote by R_1' the open region obtained from R_1 by deleting those points of the closures of $VUGH$ and $YZJI$ which are in R_1. Denote by R_2' the open region similarly obtained from R_2. The functions $\psi_1(x, y)$, $\psi_{1x}(x, y)$ and $\psi_{1y}(x, y)$ have definite finite continuous limits, the last positive, on the segment HG of l_1. The functions $\psi_2(x, y)$, $\psi_{2x}(x, y)$ and $\psi_{2y}(x, y)$ have definite finite continuous limits, the last positive, on the segment VU of l_2. From (14),

$$\lim_{\text{on } l_2} \psi_2(x, y) < M < N < \lim_{\text{on } l_1} \psi_1(x, y). \tag{15}$$

Hence there exists a function $\omega_A(y)$ such that

(a) $\omega_A(y)$ is defined and is of class C′ on s_A;

(b) $\omega_A'(y) > 0$ on s_A;

(c) if x_A is the abscissa of s_A, then $\omega_A(y) \equiv \psi_1(x_A, y)$ if (x_A, y) is in R_1', and $\omega_A(y) \equiv \psi_2(x_A, y)$ if (x_A, y) is in R_2'.

Denote by $o(s_A)$ the open, simply connected region consisting of all points belonging to characteristic-segments which lie in R and pass through s_A. The open regions $VUGH$ and $o(s_A)$ together make up an open, simply connected region: denote this region by R_A. From the proof of Lemma 1 (see footnote 2) it is clear that if R_A is identified with g, then there exists a set of open regions q_i with the properties there described and such that $q_1 \equiv o(s_A)$. Hence, by Theorem 1, there exists a function $\psi_A(x, y)$ such that

(a) $\psi_A(x, y)$ is defined and is of class C′ with respect to x and y in R_A;

(b) $\psi_A(x, y)$ satisfies (1) in R_A;

(c) $\psi_A(x, y)$ satisfies (2) in R_A;

(d) in q_1 [$\equiv o(s_A)$], $\psi_A(x, y) \equiv \omega[\varphi(x_A, x, y)]$, where $y = \varphi(x, \xi, \eta)$ is the characteristic of (1) through the point (ξ, η).

Denote by R_B the corresponding open, simply connected region made up of the open regions $YZJI$ and $o(s_B)$. There exists a function $\psi_B(x, y)$ possessing in R_B the properties corresponding to those possessed by $\psi_A(x, y)$ in R_A.

The open regions R_1', R_2', R_A and R_B together make up the region R. Each of the functions $\psi_1(x, y)$, $\psi_2(x, y)$, $\psi_A(x, y)$ and $\psi_B(x, y)$ possesses in its region of definition the properties described in the conclusion of Theorem 2, and also

$$\begin{cases} \psi_A(x, y) \equiv \psi_1(x, y) \text{ if } (x, y) \text{ is in both } R_A \text{ and } R_1'; \\ \psi_A(x, y) \equiv \psi_2(x, y) \text{ if } (x, y) \text{ is in both } R_A \text{ and } R_2'; \\ \psi_B(x, y) \equiv \psi_1(x, y) \text{ if } (x, y) \text{ is in both } R_B \text{ and } R_1'; \\ \psi_B(x, y) \equiv \psi_2(x, y) \text{ if } (x, y) \text{ is in both } R_B \text{ and } R_2'. \end{cases} \tag{16}$$

For example, if (x, y) is in both R_A and R_1',

$$\psi_A(x, y) \equiv \omega_A[\varphi(x_A, x, y)] \equiv \psi_1[x_A, \varphi(x_A, x, y)] \equiv \psi_1(x, y). \tag{17}$$

Hence the four functions $\psi_1(x, y)$, $\psi_2(x, y)$, $\psi_A(x, y)$ and $\psi_B(x, y)$ taken over their respective regions of definition R_1', R_2', R_A and R_B constitute a function $\psi(x, y)$ which possesses the properties described in the conclusion of Theorem 2.

Remark 1. It is clear that there exists a function $\psi(x, y)$ as described in Theorem 2 which possesses the additional property

(d) $L < \psi(x, y) < U$ in R,

where L and U are any preassigned finite real numbers such that $L < U$.

Remark 2. Examples show that part (c) of the hypothesis of Theorem 2 is not necessary in order that there exist a function $\psi(x, y)$ as described, but that if part (c) is omitted the remaining conditions do not imply the existence of such a function.

THEOREM 3. *Let R be an open, triply connected plane region bounded by three simple closed curves K, k_1 and k_2 such that k_1 and k_2 are contained in the open region enclosed by K. Let (x, y) be the rectangular Cartesian coördinates of a general point P of the plane for a particular coördinate system. Let $f(x, y)$ be a function such that*

(a) *$f(x, y)$ is defined and is continuous in the closure of R and is of class* C′ *with respect to y in R;*

(b) *$f_y(x, y)$ has a definite finite continuous limit on the boundary of R;*

(c) *There is a leftmost point A of k_1 through which there do not pass two distinct characteristics of* (1), *each having a limit point on K or k_2 to the right of A; there is a rightmost point B of k_1 through which there do not pass two distinct characteristics of* (1), *each having a limit point on K or k_2 to the left of B; the corresponding conditions are satisfied with reference to k_2.*

Then there exists a function $\psi(x, y)$ such that in R:

(a) *$\psi(x, y)$ is defined and is of class* C′ *with respect to x and y;*

(b) *$\psi(x, y)$ satisfies* (1);

(c) *$\psi(x, y)$ satisfies* (2).

Proof. For definiteness assume that A is as far to the left as any point on k_1 or on k_2. By a method similar to that used in the proof of Theorem 2 it can be shown that there exists a closed vertical segment t_A such that

(a) t_A is contained in an open vertical segment s_A which, including its endpoints, lies in R and is to the left of A;

(b) through each point of t_A there passes a characteristic of (1) whose left endpoint is on K and whose right endpoint is on k_1;

and that there exists a closed vertical segment t_B such that

(a) t_B is contained in an open vertical segment s_B which, including its endpoints, lies in R and is to the right of B;

(b) through each point of t_B there passes a characteristic of (1) whose left endpoint is on k_1 and whose right endpoint is either on k_2 or on K.

Suppose that through every such t_B there pass at least one characteristic whose right endpoint is on k_2 and at least one characteristic whose right endpoint is on K. Then it is possible to select a closed subsegment p_B of t_B through

whose upper endpoint passes a characteristic c_1 whose right endpoint is on k_2 and through whose lower endpoint passes a characteristic c_2 whose right endpoint is on k_2. In order that the right endpoint of a characteristic through p_B be on K it is necessary that this characteristic cross either c_1 or c_2. This is impossible. The investigation therefore reduces to the following two cases:

Case 1. *The right endpoint of every characteristic through t_B is on k_2.* It can be assumed without loss of generality that no characteristic through t_B touches K.

Denote by C a rightmost point of k_2 which satisfies part (c) of the hypothesis of the theorem. It can be shown that there exists a closed vertical segment t_C such that

(a) t_C is contained in an open vertical segment s_C which, including its endpoints, lies in R and is to the right of C;

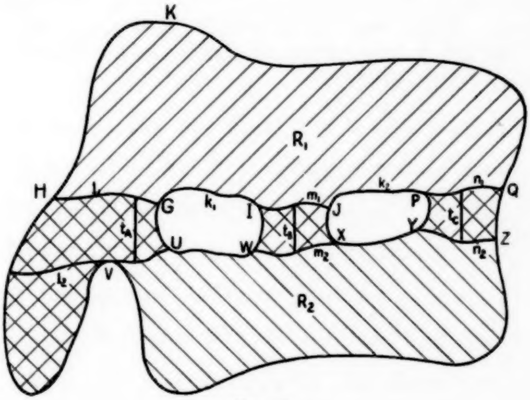

FIG. 5

(b) through each point of t_C there passes a characteristic of (1) whose left endpoint is on k_2 and whose right endpoint is on K.

Denote by l_1 and l_2 characteristics through the upper and lower endpoints, respectively, of t_A. Denote by G the leftmost limit point of l_1 on k_1, and by H the rightmost limit point of l_1 on K and to the left of G. Let U and V denote the corresponding points with reference to l_2. Assume, without loss of generality, that t_A is so small and so placed on s_A that H and V are to the left of the ordinate through t_A. Make the corresponding denotations with reference to t_B and to t_C, as indicated in Figure 5.

Denote by R_1 the open region enclosed by the curve which passes from U to W in a positive sense on k_1, from W to X on m_2, from X to Y in a positive sense on k_2, from Y to Z on n_2, from Z to V in a positive sense on K, and from V to U on l_2. Denote by R_2 the open region enclosed by the corresponding curve $PYXJIWUGHVZQP$. Each of these regions is an open, simply connected region satisfying with $f(x, y)$ the hypothesis of Theorem 1. Hence there

exist functions $\psi_1(x, y)$ and $\psi_2(x, y)$ possessing, in R_1 and R_2 respectively, the properties described in the conclusion of Theorem 1 and also satisfying (14), where M and N are any finite real numbers such that $M < N$.

Let $VUGH$ denote the open region enclosed by the curve which passes from V to U on l_2, from U to G in a positive sense on k_1, from G to H on l_1, and from H to V in a positive sense on K. Let $WXJI$ and $YZQP$ denote the corresponding open regions with reference to t_B and t_C respectively. Denote by R_1' the open region obtained from R_1 by deleting those points of the closures of these three regions which are in R_1. Denote by R_2' the open region similarly obtained from R_2.

The remainder of the proof is analogous to the proof of Theorem 2. Denote by R_A the open, simply connected region made up of $VUGH$ and $o(s_A)$; by R_B the region made up of $WXJI$ and $o(s_B)$; and by R_C the region made up of $YZQP$ and $o(s_C)$. There exist functions $\psi_A(x, y)$, $\psi_B(x, y)$ and $\psi_C(x, y)$ which possess, in R_A, R_B and R_C respectively, the properties described in the conclusion of Theorem 3 and which satisfy

$$(18)\qquad \begin{cases} \psi_A(x, y) \equiv \psi_1(x, y) \text{ if } (x, y) \text{ is in both } R_A \text{ and } R_1'; \\ \psi_A(x, y) \equiv \psi_2(x, y) \text{ if } (x, y) \text{ is in both } R_A \text{ and } R_2'; \\ \psi_B(x, y) \equiv \psi_1(x, y) \text{ if } (x, y) \text{ is in both } R_B \text{ and } R_1'; \\ \psi_B(x, y) \equiv \psi_2(x, y) \text{ if } (x, y) \text{ is in both } R_B \text{ and } R_2'; \\ \psi_C(x, y) \equiv \psi_1(x, y) \text{ if } (x, y) \text{ is in both } R_C \text{ and } R_1'; \\ \psi_C(x, y) \equiv \psi_2(x, y) \text{ if } (x, y) \text{ is in both } R_C \text{ and } R_2'. \end{cases}$$

The five functions $\psi_1(x, y)$, $\psi_2(x, y)$, $\psi_A(x, y)$, $\psi_B(x, y)$ and $\psi_C(x, y)$ taken over their respective regions of definition R_1', R_2', R_A, R_B and R_C constitute a function $\psi(x, y)$ which possesses the properties described in the conclusion of Theorem 3.

Case 2. *Case* 1 *does not occur. The right endpoint of every characteristic through t_B is on K.* It can be assumed without loss of generality that no characteristic through t_B touches k_2.

Denote by m_1 and m_2 characteristics through the upper and lower endpoints, respectively, of t_B. Denote by I the rightmost limit point of m_1 on k_1, and by J the leftmost limit point of m_1 on K and to the right of I. Let W and X denote the corresponding points with reference to m_2. Assume, without loss of generality, that t_B is so small and so placed on s_B that J and X are to the right of the ordinate through t_B.

Denote by R_1 the open region enclosed by the curve which passes from U to W in a positive sense on k_1, from W to X on m_2, from X to V in a positive sense on K, and from V to U on l_2. Denote by R_2 the open region enclosed by the corresponding curve $IWUGHVXJI$. Each of these regions is an open, at most doubly connected region satisfying with $f(x, y)$ the hypothesis either of

Theorem 1 or of Theorem 2. Hence there exist functions $\psi_1(x, y)$ and $\psi_2(x, y)$ possessing, in R_1 and R_2 respectively, the properties described in the conclusion of Theorem 1 and also satisfying (14), where M and N are any finite real numbers such that $M < N$.

Let $WXJI$ denote the open region enclosed by the curve which passes from W to X on m_2, from X to J in a positive sense on K, from J to I on m_1, and from I to W in a positive sense on k_1. Suppose that every segment t_B as described above is such that the region $WXJI$ which it determines is doubly connected. Then the characteristic-segment PQ through the midpoint of t_B corresponding to the segment IJ of the characteristic m_1 cuts the region $WXJI$ into two distinct regions, each of which is doubly connected. The region R is therefore quadruply connected. This contradicts the hypothesis of the theorem.

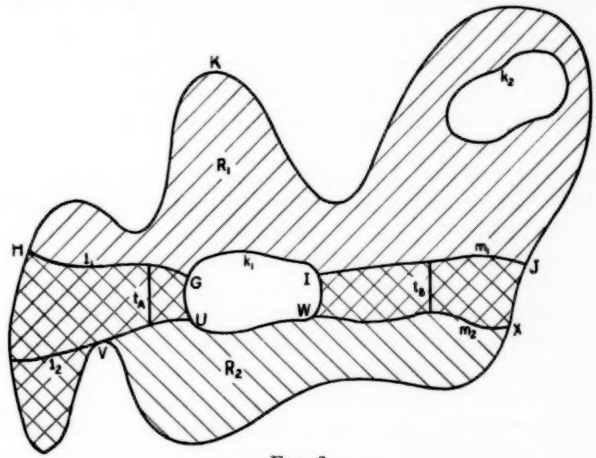

FIG. 6

Hence it can be assumed without loss of generality that the segment t_B is such that the region $WXJI$ is simply connected.

The remainder of the proof is analogous to the proof for Case 1. Define the open regions R_1', R_2', R_A and R_B. There exist functions $\psi_A(x, y)$ and $\psi_B(x, y)$ which possess, in R_A and R_B respectively, the properties described in the conclusion of Theorem 3 and which satisfy (16). The four functions $\psi_1(x, y)$, $\psi_2(x, y)$, $\psi_A(x, y)$ and $\psi_B(x, y)$ taken over their respective regions of definition R_1', R_2', R_A and R_B constitute a function $\psi(x, y)$ which possesses the properties described in the conclusion of Theorem 3.

Remark 1. It is clear that there exists a function $\psi(x, y)$ as described in Theorem 3 which possesses the additional property

(d) $L < \psi(x, y) < U$ in R,

where L and U are any preassigned finite real numbers such that $L < U$.

Remark 2. A theorem analogous to Theorems 1, 2 and 3 can be stated for

regions of order of connectivity m, where m is any positive integer. In the manner indicated in the proofs of Theorems 2 and 3, the proof of the existence of a solution of (1) having the desired properties in a region of order of connectivity m (>1) can be made to depend upon the existence of such solutions in regions of order of connectivity less than m.

Bibliography

1. E. Kamke, *Zur Theorie der Differentialgleichungen*, Mathematische Annalen, vol. 99(1928), pp. 602–614.
2. T. Ważewski, *Sur l'équation $p + Qq = 0$*, Mathematica, vol. 8(1934), pp. 103–116.
3. H. Whitney, *Analytic extensions of differentiable functions defined in closed sets*, Transactions of the American Mathematical Society, vol. 36(1934), pp. 63 ff.

The University of Alabama.

A CONVERSE THEOREM CONCERNING THE DIAMETRAL LOCUS OF AN ALGEBRAIC CURVE

By Jesse Douglas

1. **Introduction.** If a conic be cut by any system of parallel secants, the locus of the midpoint of the two intersections with each secant is a straight line, called a diameter of the conic. This diameter is the polar with respect to the conic of the common infinite point of the parallel secants.

It was proved by Sir Isaac Newton[1] that the preceding elementary property of conics extends to an algebraic curve A of any degree n. The intersection of A with any straight line l is a system of n real or imaginary points; let G denote their centroid. Then the locus of G as l moves parallel to itself is a straight line d, which may be called a *diameter* of A. The line d is, in fact, the linear polar with respect to A of the fixed infinite point of the system of parallel secants.

If any n curve-elements $\gamma_1, \gamma_2, \cdots, \gamma_n$ are cut by a system of parallel secants, we naturally define the corresponding "diametral locus" as the locus of the centroid G of the respective intersection-points $p_1, p_2, \cdots, p_n$ of the given curve-elements with an arbitrary secant l of the given parallel system. Suppose that this diametral locus is a straight line *for every system of parallel secants.* Then we shall prove in this paper (Theorem II) that the curve-elements γ_i must belong to the same algebraic curve of degree n, possibly a reducible one.

Let $\sigma_1, \sigma_2, \cdots, \sigma_n$ denote n hypersurface-elements in Euclidean space of any number $m + 1$ of dimensions. If the diametral locus of these hypersurface-elements relative to an arbitrary system of parallel secants is a hyperplane, then the elements must belong to an algebraic hypersurface of degree n, possibly reducible. This theorem (III), the natural analogue in higher dimensions of the one stated in the preceding paragraph, is proved in §6.

We conclude (§7) with a discussion of the conditions on the curve- and hypersurface-elements γ_i, σ_i under which our results are function-theoretically valid.

2. **Relation to other literature.** Our results include as very special cases certain converse theorems concerning polynomials, $y = P(x)$, recently given by Howard Levi.[2] This author states his result in an analytic form, which can be seen to amount geometrically to this: If the diametral locus, relative to an arbitrary direction, determined by n elements of an entire function $y = f(x)$ is a vertical straight line, then the entire function must be a polynomial. This

Received September 11, 1939.

[1] See Salmon, *Higher Plane Curves*, Dublin, 1873, p. 109.

[2] *On the values assumed by polynomials,* Bull. Amer. Math. Soc., vol. 45(1939), pp. 570–575.

is, obviously, an immediate corollary of the theorem stated in the previous section, since the only entire functions which are algebraic are the polynomials. In fact, Levi's theorem already follows from a restricted and more easily proved form (Theorem I) of our main theorem (II), the restriction consisting in the imposition of an invariable direction on the diameter.

In the case of hypersurfaces, the simplification achieved by taking the geometric point of view is even more striking.[3]

Some of our calculations, particularly in §§3, 4, are similar to those of Levi's interesting paper, but take a more symmetric and elegant form because of our more natural mode of treatment.

3. **Formulas.** We first establish the basic formulas which underly all our calculations and proofs.

Consider n curve-elements γ_i :

$$y = f_i(x) \qquad (i = 1, 2, \cdots, n). \tag{3.1}$$

Let the straight line

$$y = kx + c \tag{3.2}$$

intersect these curve-elements in the respective points p_i of coördinates (x_i, y_i). Then x_i is a function of k, c which obeys the equation

$$f_i(x_i) = kx_i + c. \tag{3.3}$$

By differentiation with respect to k and with respect to c:

$$f_i'(x_i) \frac{\partial x_i}{\partial k} = x_i + k \frac{\partial x_i}{\partial k},$$

$$f_i'(x_i) \frac{\partial x_i}{\partial c} = k \frac{\partial x_i}{\partial c} + 1.$$

Eliminating $f_i'(x_i)$, we find that $x_i(k, c)$ obeys the partial differential equation

$$\frac{\partial x_i}{\partial k} = x_i \frac{\partial x_i}{\partial c}, \tag{3.4}$$

or

$$\frac{1}{x_i} \frac{\partial x_i}{\partial k} = \frac{\partial x_i}{\partial c}. \tag{3.4'}$$

The n quantities x_i are the roots of the algebraic equation

$$(x - x_1)(x - x_2) \cdots (x - x_n) = 0,$$

or

$$x^n - P_1x^{n-1} + P_2x^{n-2} + \cdots + (-1)^{n-i}P_{n-i}x^i + \cdots + (-1)^nP_n = 0, \tag{3.5}$$

[3] Cf. Levi, loc. cit., §2.

where $P_1, P_2, \cdots, P_n$ are the elementary symmetric functions of the x_i's. Hence the P_i's are, like the x_i's, also functions of k, c, and we have next to find what relations among the P_i's are implied by the partial differential equation (3.4) obeyed by each x_i.

First, we obtain by addition of (3.4′) for $i = 1, 2, \cdots, n$:

$$\frac{1}{P_n}\frac{\partial P_n}{\partial k} = \frac{\partial P_1}{\partial c}. \tag{3.6}$$

Next, let us differentiate

$$P_i = \sum x_{\alpha_1} x_{\alpha_2} \cdots x_{\alpha_i}$$

with respect to k. We find, after use of (3.4),

$$\frac{\partial P_i}{\partial k} = \sum x_{\alpha_1} x_{\alpha_2} \cdots x_{\alpha_i} \frac{\partial x_\beta}{\partial c}. \tag{3.7}$$

Here the sum $\sum$ refers to every possible combination of indices $\alpha_1, \alpha_2, \cdots, \alpha_i, \beta$ (each ranging from 1 to n) where the α's are distinct and *β is equal to one of the α's.*

By differentiation with respect to c of

$$P_{i+1} = \sum x_{\alpha_1} x_{\alpha_2} \cdots x_{\alpha_i} x_{\alpha_{i+1}},$$

we find

$$\frac{\partial P_{i+1}}{\partial c} = (i + 1) \sum{}' x_{\alpha_1} x_{\alpha_2} \cdots x_{\alpha_i} \frac{\partial x_\gamma}{\partial c}, \tag{3.8}$$

where the sum $\sum'$ refers to every possible combination of indices $\alpha_1, \alpha_2, \cdots, \alpha_i, \gamma$ in which the α's are distinct and *γ is distinct from each α.* Hence, adding (3.7) to (3.8) after dividing the latter by $i + 1$, we get

$$\frac{\partial P_i}{\partial k} + \frac{1}{i+1}\frac{\partial P_{i+1}}{\partial c} = \sum{}'' x_{\alpha_1} x_{\alpha_2} \cdots x_{\alpha_i} \frac{\partial x_\delta}{\partial c},$$

where in the summation $\sum''$ every combination of indices is allowed to occur (once and only once) in which the α's are distinct while the index *δ is unrestricted.* Thus, the right member of the last relation can be written

$$\sum x_{\alpha_1} x_{\alpha_2} \cdots x_{\alpha_i} \cdot \sum \frac{\partial x_\delta}{\partial c},$$

so that this relation becomes

$$\frac{\partial P_i}{\partial k} + \frac{1}{i+1}\frac{\partial P_{i+1}}{\partial c} = P_i \frac{\partial P_1}{\partial c}. \tag{3.9}$$

4. **Particular algebraic curves.** Let us now formulate our hypothesis concerning the rectilinear character of the diametral locus of the n curve-elements γ_i in the following restricted form: The diametral locus is not only a straight

line for every fixed direction k, but a straight line of *invariable direction* (independent of k). Evidently, we may take this direction to be the vertical one, x = const., without loss of any generality.

Then, G denoting the centroid of the n intersection-points p_i, we have

$$x_G = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{P_1}{n};$$

so that the condition: x_G = const. for fixed k, is equivalent to: P_1 = const. for fixed k; i.e., P_1 is a function of k alone, not of c, or

$$\frac{\partial P_1}{\partial c} = 0. \tag{4.1}$$

Thereby, (3.6) and (3.9) give respectively

$$\frac{\partial P_n}{\partial k} = 0, \tag{4.2}$$

$$\frac{\partial P_i}{\partial k} = -\frac{1}{i+1}\frac{\partial P_{i+1}}{\partial c}. \tag{4.3}$$

By (4.2),

$$P_n = \varphi(c), \tag{4.4}$$

a function of c alone. Substitute this in (4.3) with $i = n - 1$; then

$$\frac{\partial P_{n-1}}{\partial k} = -\frac{1}{n}\varphi'(c);$$

whence, by integration,

$$P_{n-1} = -\frac{1}{n}\varphi'(c)k + \varphi_1(c). \tag{4.5}$$

In (4.3) put $i = n - 2$, substitute (4.5), and integrate; then

$$P_{n-2} = \frac{1}{n(n-1)}\varphi''(c)\frac{k^2}{2} - \frac{1}{n-1}\varphi_1'(c)k + \varphi_2(c). \tag{4.6}$$

Continuing in this way, we obtain generally

$$\begin{aligned} P_{n-i} = N_i\varphi^{(i)}(c)k^i + N_{i-1}\varphi_1^{(i-1)}(c)k^{i-1} + \cdots \\ + N_{i-j}\varphi_j^{(i-j)}(c)k^{i-j} + \cdots + \varphi_i(c). \end{aligned} \tag{4.7}$$

Here the N's denote purely numerical coefficients, while the indices in parentheses denote an order of differentiation.

If we allow i to be $n - 1$ in (4.7), and then apply the fact that P_1 is a function of k alone, it follows that $\varphi^{(n-1)}(c), \varphi_1^{(n-2)}(c), \cdots, \varphi_j^{(n-j-1)}(c), \cdots, \varphi_{n-1}(c)$ are constants. Accordingly, $\varphi(c), \varphi_1(c), \cdots, \varphi_j(c), \cdots, \varphi_{n-1}(c)$ are polynomials

in c, of respective degrees $n - 1, n - 2, \cdots, n - j - 1, \cdots, 0$ at most (if, indeed, they do not vanish identically).[4]

Using (4.7), let us form the general term $P_{n-i}x^i$ of (3.5), and then write by (3.2): $kx = y - c$; we get a result of the form

$$\begin{aligned} P_{n-i}x^i = \psi_0(c)(y - c)^i + \psi_1(c)x(y - c)^{i-1} + \cdots \\ + \psi_j(c)x^j(y - c)^{i-j} + \cdots + \psi_i(c)x^i, \end{aligned} \tag{4.8}$$

where each $\psi_j(c)$ is a certain polynomial in c. Consequently,

$$P_{n-i}x^i = Q_i(x, y, c), \tag{4.8'}$$

where Q_i is a polynomial in x, y, c of degree i in x, y—unless, indeed, $Q_i \equiv 0$.

If we substitute this in (3.5), it follows that each intersection-point p_i, or (x_i, y_i), of the line $y = kx + c$ with the curve-elements γ_i obeys the equation in x, y:

$$\begin{aligned} x^n - Q_{n-1}(x, y, c) + Q_{n-2}(x, y, c) + \cdots \\ + (-1)^{n-i}Q_i(x, y, c) + \cdots + (-1)^n\varphi(c) = 0. \end{aligned} \tag{4.9}$$

This equation must hold for every value of c that belongs as y-intercept to a straight line $y = kx + c$ passing through the point (x, y), provided that this straight line takes part in the construction of any of our diametral loci. Let us assume, as hypothesis, the existence of an angular pencil of lines with vertex at $(0, c_0)$ such that every line in the angular pencil intersects every curve-element γ_i in a single point.[5] Then in (4.9) we may put $c = c_0$ and have finally, in a condensed notation,

$$x^n + R_{n-1}(x, y) = 0, \tag{4.10}$$

where R_{n-1} is a polynomial in x, y of degree $n - 1$ at most—or, possibly, $R_{n-1}(x, y) \equiv 0$.

The equation (4.10) applies to the portions of the elements γ_i which are covered, as described, by the angular pencil of lines with vertex at $(0, c_0)$. This point may be considered as simply any one of the plane, since the axes of x, y are subject to arbitrary translation without changing anything essential. By varying the position of the vertex-point, the portions of the γ_i's to which our reasoning applies can be made to overlap so as wholly to cover the γ_i's. Hence,

[4] Here and at many similar places hereafter, the phraseology could be abbreviated by the use of a special term, like "polynomial of grade n", to denote any expression of the form

$$a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n,$$

regardless of how many of the leading coefficients are zero. Thus a polynomial of grade n means one which is of degree $\leqq n$ or which vanishes identically. The same terminology can be applied to polynomials in several independent variables.

[5] This assumption is justified by the discussion of §7.

(4.10) being algebraic, our reasoning shows that this one equation applies to the curve-elements γ_i in their entirety.

(4.10) is simply the general algebraic equation of n-th degree in x, y restricted by the condition that the group of terms of n-th degree reduces to x^n. This corresponds to the fact that $x = \text{const.}$ is the invariable diametral direction; if this invariable direction is $ax + by = \text{const.}$, then, as we easily see by affine transformation, the corresponding form of equation (4.10) is

$$(4.11) \qquad (ax + by)^n + R_{n-1}(x, y) = 0.$$

Thus, we have our first theorem, a preliminary form of the main one (Theorem II) to be proved later, and differing from the latter only in the requirement of a fixed diametral direction.

THEOREM I. *If the diametral locus determined by n curve-elements γ_i with respect to an arbitrary system of parallel secants is a straight line of invariable direction, $ax + by = \text{const.}$, then the elements γ_i belong to an algebraic curve of n-th degree having the particular form* (4.11). *This algebraic curve may be reducible.*

It may be observed, as an easy corollary, that if the diametral line is absolutely fixed: $ax + by + C = 0$, regardless of the particular system of parallel secants, then (4.11) is further specialized to the form

$$(4.12) \qquad (ax + by)^n + nC(ax + by)^{n-1} + R_{n-2}(x, y) = 0,$$

where R_{n-2} (if not $\equiv 0$) is a polynomial of degree $n - 2$ at most in x, y.

5. **General algebraic curves.** We now give up the condition of invariable direction for our diametral locus, and require only that this locus be a straight line for every fixed direction k. This is expressed by

$$(5.1) \qquad y_G = g(k)x_G + h(k).$$

Since (x_G, y_G) lies upon the line (3.2),

$$(5.2) \qquad y_G = kx_G + c;$$

therefore, by solution for x_G of the last two equations,

$$(5.3) \qquad x_G = \frac{c - h(k)}{g(k) - k},$$

unless $g(k) = k$. But $g(k) = k$ would imply the parallelism, and therefore the identity, of the lines (5.1), (5.2), since they have the point (x_G, y_G) in common. Then the γ_i would have to be rectilinear elements of the diameter (5.1), and could therefore be regarded as part of a reducible algebraic curve of degree n, namely, the diameter taken n-fold. Accordingly, we may lay this case aside as trivial, and suppose $g(k) \not\equiv k$.

Since $x_G = P_1/n$, formula (5.3) leads to one of the form

$$(5.4) \qquad P_1 = \lambda(k)c + \mu(k),$$

stating that P_1 is a linear function with respect to c. We shall draw our inferences from this implication of the diametral condition.

Substituting (5.4) in (3.6), we get

$$\frac{1}{P_n}\frac{\partial P_n}{\partial k} = \lambda(k),$$

whence, by integration,

$$P_n = \varphi(c)e^{\int \lambda(k)\,dk}. \tag{5.5}$$

Further, by (5.4), (3.9), we have

$$\frac{\partial P_i}{\partial k} - \lambda(k)P_i = -\frac{1}{i+1}\frac{\partial P_{i+1}}{\partial c}. \tag{5.6}$$

Let us define

$$\bar{P}_i = e^{-\int \lambda(k)\,dk}P_i\ ; \tag{5.7}$$

then, multiplying (5.6) by $e^{-\int \lambda(k)\,dk}$, we have

$$\frac{\partial \bar{P}_i}{\partial k} = -\frac{1}{i+1}\frac{\partial \bar{P}_{i+1}}{\partial c}. \tag{5.8}$$

By (5.5, 7),

$$\bar{P}_n = \varphi(c). \tag{5.9}$$

Comparing (5.8, 9) with (4.3, 4), we see that exactly the same relations now hold for $\bar{P}_i$ as previously held for P_i. It follows then, by identical calculations, that, like (4.7),

$$\begin{aligned}\bar{P}_{n-i} = N_i\varphi^{(i)}(c)k^i + N_{i-1}\varphi_1^{(i-1)}(c)k^{i-1} + \cdots \\ + N_{i-j}\varphi_j^{(i-j)}(c)k^{i-j} + \cdots + \varphi_i(c).\end{aligned} \tag{5.10}$$

There is, however, this one difference: that now if we put $i = n - 1$, giving $\bar{P}_1$, this quantity is by (5.4, 7) a linear function of c instead of a function of k alone, as before. Therefore $\varphi^{(n-1)}(c)$, $\varphi_1^{(n-2)}(c)$, $\cdots$, $\varphi_j^{(n-j-1)}(c)$, $\cdots$, $\varphi_{n-1}(c)$ are linear functions of c, and consequently the polynomials $\varphi(c)$, $\varphi_1(c)$, $\cdots$, $\varphi_j(c)$, $\cdots$, $\varphi_{n-1}(c)$ have the respective degrees n, $n - 1$, $\cdots$, $n - j$, $\cdots$, 1 at most.

Multiply (5.10) by x^i, and write $kx = y - c$; then

$$\begin{aligned}\bar{P}_{n-i}x^i = \psi_0(c)(y-c)^i + \psi_1(c)x(y-c)^{i-1} + \cdots \\ + \psi_j(c)x^j(y-c)^{i-j} + \cdots + \psi_i(c)x^i,\end{aligned} \tag{5.11}$$

each $\psi_j(c)$ being a polynomial of degree $(n - j) - (i - j) = n - i$ at most in c. It follows that

$$\bar{P}_{n-i}x^i = Q_i(x, y, c), \tag{5.11$'$}$$

where Q_i is a polynomial in x, y, c of degree i in x, y and of degree n at most in c—unless, perhaps, $Q_i \equiv 0$.[6]

The equation (3.5), multiplied by

$$\psi(k) = e^{-\int \lambda(k)\, dk}, \tag{5.12}$$

is, according to (5.7),

$$\psi(k)x^n - \overline{P}_1 x^{n-1} + \cdots + (-1)^{n-i}\overline{P}_{n-i}x^i + \cdots + (-1)^n\overline{P}_n = 0.$$

By (5.11′), and the substitution $k = (y - c)/x$, this becomes

$$\psi\left(\frac{y-c}{x}\right)x^n - Q_{n-1}(x, y, c) + \cdots + (-1)^{n-i}Q_i(x, y, c) + \cdots + (-1)^n\varphi(c) = 0. \tag{5.13}$$

We may rewrite this as

$$H_n(x, y - c) + R_{n-1}(x, y, c) = 0. \tag{5.14}$$

Here

$$H_n(x, y - c) = x^n\psi\left(\frac{y-c}{x}\right) \tag{5.15}$$

denotes a *homogeneous function* of degree n in x, $y - c$, which *does not vanish identically* because of the exponential form (5.11) of ψ. R_{n-1} denotes a polynomial in x, y, c of degree $n - 1$ at most in x, y and degree n at most in c—or, possibly, $R_{n-1} \equiv 0$.[7]

By its derivation, the relation (5.14) must be obeyed by the coördinates (x, y) of every point on every curve-element γ_i and for an interval of values of c. In other words, regardless of the value of c, (5.14) must define the same n function-elements $y = f_i(x)$. If we interpret (x, y, c) as coördinates in a three-space, (5.14) defines a surface S, and for any fixed value c_0 of c, the curve or curves $y = f_i(x)$ defined implicitly by (5.14) are found by cutting S with the plane $c = c_0$ and projecting orthogonally on the xy-plane. Since these curves γ_i, or $y = f_i(x)$, are the same regardless of the value of c_0, it follows that the surface S must be a right cylinder based on the curves γ_i. Hence *the left member of the equation* (5.14) *of S must, in reality, be independent of c*; i.e.,

$$H_n(x, y - c) \equiv - R_{n-1}(x, y, c) + F(x, y). \tag{5.16}$$

Here write $c = y - z$; then

$$H_n(x, z) \equiv S(x, y, z) + F(x, y),$$

where, since R_{n-1} is a polynomial in c of degree n at most, S is a polynomial in z of degree n at most. Hence this is true of H_n:

$$H_n(x, z) \equiv \sum_{i=0}^{n} \xi_i(x)z^i.$$

[6] In fact, the total degree of Q_i in x, y, c is n at most.

[7] In fact, the total degree of R_{n-1} in x, y, c is n at most.

But since H_n is homogeneous of degree n jointly in x, z, it follows easily that $\xi_i(x) \equiv C_i x^{n-i}$ (C_i a constant); therefore H_n, not vanishing identically, is a *homogeneous polynomial of degree n in x, z.*

Consequently, giving any particular value to c in (5.14), as we may,[8] we see that this equation has the form of the *general algebraic equation of the n-th degree in x, y:*

$$K_n(x, y) + T_{n-1}(x, y) = 0, \tag{5.17}$$

K_n denoting a homogeneous polynomial of the n-th degree (certainly not vanishing identically), while T_{n-1} is a polynomial of degree $n - 1$ at most—unless, perhaps, $T_{n-1} \equiv 0$.

Thus we have proved

THEOREM II. *If the diametral locus determined by n given curve-elements γ_i relative to an arbitrary system of parallel secants is a straight line, then the elements γ_i belong to the same algebraic curve of degree n. This curve may be reducible.*

6. **Characterization of algebraic hypersurfaces by the diametral property.** The preceding theory for curves extends in the most natural way to hypersurfaces.

Let x, y_α ($\alpha = 1, 2, \cdots, m$) be Cartesian coördinates in $(m + 1)$-dimensional Euclidean space. Any n given hypersurface-elements σ_i will be represented by equations of the form

$$x = f_i(y_\alpha) \qquad (\alpha = 1, \cdots, m;\, i = 1, \cdots, n). \tag{6.1}$$

Any straight line is represented by a system of equations of the form

$$y_\alpha = k_\alpha x + c_\alpha \qquad (\alpha = 1, \cdots, m). \tag{6.2}$$

The intersection-point p_i (x_i, y_i) of this line and the hypersurface-element σ_i is defined by substituting (6.2) in (6.1):

$$x_i = f_i(k_\alpha x + c_\alpha). \tag{6.3}$$

By differentiation with respect to k_α and c_α,

$$\frac{\partial x_i}{\partial k_\alpha} = x_i f_{i\alpha}(k_\beta x_i + c_\beta), \qquad \frac{\partial x_i}{\partial c_\alpha} = f_{i\alpha}(k_\beta x_i + c_\beta) \qquad (\beta = 1, \cdots, m),$$

where $f_{i\alpha} \equiv \partial f_i / \partial y_\alpha$. Therefore

$$\frac{\partial x_i}{\partial k_\alpha} = x_i \frac{\partial x_i}{\partial c_\alpha} \quad \text{or} \quad \frac{1}{x_i}\frac{\partial x_i}{\partial k_\alpha} = \frac{\partial x_i}{\partial c_\alpha}, \tag{6.4}$$

for all values of i and α.

By the similarity of (6.4) to (3.4, 4′), we obtain in precisely the same way as in §3 formulas analogous to (3.6, 9), namely:

$$\frac{1}{P_n}\frac{\partial P_n}{\partial k_\alpha} = \frac{\partial P_1}{\partial c_\alpha}, \tag{6.5}$$

[8] Cf. the discussion following (4.9), (4.10).

$$\frac{\partial P_i}{\partial k_\alpha} + \frac{1}{i+1}\frac{\partial P_{i+1}}{\partial c_\alpha} = P_i\frac{\partial P_1}{\partial c_\alpha}. \tag{6.6}$$

Omitting consideration of any restricted case, we pass at once to the hypothesis that the locus of G, centroid of the n intersection-points p_i, is a hyperplane for every system of parallel secants, i.e., secants for which the k_α's remain fixed while the c_α's vary. This condition is expressed by

$$x_G = \sum_\alpha g_\alpha(k)y_{\alpha G} + h(k), \tag{6.7}$$

where, as always hereafter, k denotes the system (k_α).

By (6.2),

$$y_{\alpha G} = k_\alpha x_G + c_\alpha,$$

and combining this with (6.7), we find that x_G is a linear function of the c_α's:

$$x_G = \frac{\sum_\alpha g_\alpha(k)c_\alpha + h(k)}{1 - \sum_\alpha g_\alpha(k)k_\alpha}. \tag{6.8}$$

The only proviso is that $1 - \sum_\alpha g_\alpha(k)k_\alpha \neq 0$; but the contrary possibility can be dismissed by an argument like that given in connection with (5.3), for it would imply that all the elements σ_i lie in the same hyperplane, the diametral one.

Because $x_G = P_1/n$, we may express (6.8) in the form

$$P_1 = \sum_\alpha \lambda_\alpha(k)c_\alpha + \mu(k). \tag{6.9}$$

By (6.5), we then have

$$\frac{\partial \log P_n}{\partial k_\alpha} = \lambda_\alpha(k). \tag{6.10}$$

It follows that $\sum_\alpha \lambda_\alpha(k)\, dk_\alpha$ is an exact differential; its integral (after any fixed choice of the constant of integration) is a definite function of the k_α's. Then, by integration of (6.10),

$$P_n = \varphi(c)\exp\left[\int \sum_\alpha \lambda_\alpha(k)\, dk_\alpha\right], \tag{6.11}$$

where c denotes the system (c_α).

Substituting (6.9) in (6.6), we find

$$\frac{\partial P_i}{\partial k_\alpha} - \lambda_\alpha(k)P_i = -\frac{1}{i+1}\frac{\partial P_{i+1}}{\partial c_\alpha}. \tag{6.12}$$

Let us define

$$\bar{P}_i = P_i \exp\left[-\int \sum_\alpha \lambda_\alpha(k)\, dk_\alpha\right]; \tag{6.13}$$

then multiplying (6.12) by $\exp\left[-\int \sum_\alpha \lambda_\alpha(k)\, dk_\alpha\right]$, we obtain

$$\frac{\partial \bar{P}_i}{\partial k_\alpha} = -\frac{1}{i+1}\frac{\partial \bar{P}_{i+1}}{\partial c_\alpha}. \tag{6.14}$$

By (6.11), (6.13)

$$\bar{P}_n = \varphi(c). \tag{6.15}$$

Formulas (6.14), (6.15) are the same, except for the subscript α, as (5.8), (5.9), or as (4.3), (4.4). We derive therefrom by quite analogous reasoning:

$$\bar{P}_{n-1} = -\frac{1}{n}\sum_\alpha \varphi_\alpha(c)k_\alpha + \varphi^{(1)}(c), \tag{6.16}$$

$$\bar{P}_{n-2} = \frac{1}{n(n-1)}\cdot\frac{1}{2}\sum_{\alpha\beta}\varphi_{\alpha\beta}(c)k_\alpha k_\beta - \frac{1}{n-1}\sum_\alpha \varphi_\alpha^{(1)}(c)k_\alpha + \varphi^{(2)}(c), \tag{6.17}$$

and so on. Here $\varphi_\alpha(c) \equiv \partial\varphi/\partial c_\alpha$, $\varphi_{\alpha\beta}(c) \equiv \partial^2\varphi/\partial c_\alpha\partial c_\beta$, etc. In general, by use also of the linear character (6.9) of P_1 with respect to the c's, $\bar{P}_{n-i}$ is seen to be a polynomial in the system of variables c, k of total degree n at most and of degree i at most in the k's—unless $\bar{P}_{n-i} \equiv 0$.

If we substitute $k_\alpha = (y_\alpha - c_\alpha)/x$, it follows then that

$$\bar{P}_{n-i}x^i = Q_i(x, y_\alpha, c_\alpha), \tag{6.18}$$

where Q_i is a polynomial in the variables x, y_α, c_α of total degree n at most, and of degree i in x, y_α—or, possibly, $Q_i \equiv 0$.

Let us multiply (3.5) by the non-identically vanishing function

$$\psi(k) = \exp\left[-\int \sum_\alpha \lambda_\alpha(k)\, dk_\alpha\right]; \tag{6.19}$$

then it becomes by (6.13):

$$\psi(k)x^n - \bar{P}_1x^{n-1} + \cdots + (-1)^{n-i}\bar{P}_{n-i}x^i + \cdots + (-1)^n\bar{P}_n = 0. \tag{6.20}$$

By (6.18), (6.15), and the substitution $k_\alpha = (y_\alpha - c_\alpha)/x$, this takes the form

$$\psi\left(\frac{y_\alpha - c_\alpha}{x}\right)x^n - Q_{n-1}(x, y_\alpha, c_\alpha)$$
$$+ \cdots + (-1)^{n-i}Q_i(x, y_\alpha, c_\alpha) + \cdots + (-1)^n\varphi(c_\alpha) = 0,$$

or, condensing, we have

$$H_n(x, y_\alpha - c_\alpha) + R_{n-1}(x, y_\alpha, c_\alpha) = 0. \tag{6.21}$$

Here

$$H_n(x, y_\alpha - c_\alpha) = x^n\psi\left(\frac{y_\alpha - c_\alpha}{x}\right) \tag{6.22}$$

is a homogeneous function of degree n in x, $y_\alpha - c_\alpha$, which does not vanish identically, on account of the exponential form (6.19) of $\psi(k)$. R_{n-1} is a polynomial in x, y_α, c_α of total degree n at most, and of degree $n - 1$ at most in x, y_α—or, perhaps, $R_{n-1} \equiv 0$.

It follows by reasoning exactly like that given for (5.14) that $H_n(x, y_\alpha - c_\alpha)$ is, in fact, a *homogeneous polynomial* of degree n in x, $y_\alpha - c_\alpha$. Giving then to (c_α) in (6.21) any permissible particular values $(c_\alpha^{(0)})$,[9] we have finally as an equation obeyed by all the hypersurface-elements σ_i:

$$K_n(x, y_\alpha) + S_{n-1}(x, y_\alpha) = 0, \tag{6.23}$$

where K_n denotes a homogeneous polynomial of degree n in x, y_α (certainly not vanishing identically), while S_{n-1} (if not identically zero) is a polynomial of degree $n - 1$ at most in x, y_α.

This completes the proof of

THEOREM III. *If the diametral locus of n given hypersurface-elements σ_i relative to an arbitrary system of parallel secants is a hyperplane, then these elements belong to the same algebraic hypersurface of degree n (which may be reducible).*

7. **Conditions of validity.** By going back over our calculations, it is easy to determine sufficient conditions under which our results are function-theoretically valid.

First, the *real domain*. The formulas (5.10), which for $i = 0, 1, \cdots, n - 1$ give $\bar{P}_n, \bar{P}_{n-1}, \cdots, \bar{P}_1$, involve the derivatives of $\varphi(c)$ up to the $(n - 1)$-th order, of $\varphi_1(c)$ up to the $(n - 2)$-th order, $\cdots$, of $\varphi_j(c)$ up to the order $n - 1 - j$; hence the existence of all these derivatives must be provided for. Accordingly, we suppose that $\partial^{j-1}P_j/\partial c^{j-1}$ exists for $j = n, n - 1, \cdots, 1$. To enable us to pass between P_i and $\bar{P}_i$, the function $\lambda(k)$ in (5.7) should be continuous; i.e., $\partial P_1/\partial c$ $(= \lambda(k)$ by (5.4)) should be continuous. Also, the derivatives $\partial P_i/\partial k$, $\partial P_{i+1}/\partial c$ appearing in (5.6) must be existent.

Evidently, all these conditions will be fulfilled if each $x_i(k, c)$ is a function of class $C^{(n-1)}$. Since the functions $x_i(k, c)$ are defined implicitly by the equations (3.3), they will be of class $C^{(n-1)}$ if this is true of the functions $y = f_i(x)$, i.e., of the curve-elements γ_i. We suppose these elements to be cut *non-tangentially* by an initial line $y = k_0 x + c_0$, and restrict k, c to properly delimited intervals about k_0, c_0.

A similar argument and conclusion apply to the case of hypersurface-elements σ_i.

In the *complex domain*, the coördinates x, y_α and the line coefficients k_α, c_α are complex variables, which we suppose restricted to suitable regions in their respective complex planes. These regions may be taken to be sufficiently small circles about $x_i^{(0)}$, $y_{\alpha i}^{(0)}$, $k_\alpha^{(0)}$, $c_\alpha^{(0)}$, where $y_\alpha = k_\alpha^{(0)} x + c_\alpha^{(0)}$ is an initial straight line

[9] Cf. the discussion following (4.9), (4.10).

intersecting the hypersurface-element σ_i non-tangentially at $(x_i^{(0)}, y_{\alpha i}^{(0)})$ for $i = 1, 2, \cdots, n$. The elements σ_i are then naturally assumed to be *analytic*.

However, the argument certainly seems appealing that to demand that n curve-elements γ_i give a diametral locus as regular as a straight line *for every possible fixed direction* k is likely to force these curve-elements to have a high degree of regularity, without it being necessary to stipulate such a condition as hypothesis.

Without engaging upon the discussion of this question in full, an idea of what may be expected in this direction may be had by the following consideration of a particular form of the case $n = 2$.

Let $y = f(x)$ be merely a *continuous* curve, defined in an interval that we may take to be $(-1, 1)$ without loss of generality. Let c denote any point of this interval; then $[c + h, f(c + h)]$ and $[c - h, f(c - h)]$ represent in the most general way any two points P_h, Q_h of $y = f(x)$ such that the midpoint of the chord P_hQ_h lies on the vertical line $x = c$. The only proviso is that $c - h$, $c + h$ shall both belong to the interval $(-1, 1)$.

Suppose that as h varies the chord P_hQ_h always remains parallel to itself; this is our diametral condition applied to the curve $y = f(x)$. Analytically, this condition is represented by the formula

$$\frac{f(c + h) - f(c - h)}{2h} = \varphi(c), \tag{7.1}$$

where $\varphi(c)$ is supposed merely to *exist* as a one-valued function (its continuity, not necessary for our purposes, follows from (7.1) itself, since f is continuous).

We rewrite (7.1) as

$$f(c + h) = f(c - h) + 2h\varphi(c), \tag{7.2}$$

and then form the second difference of this relation with respect to h; i.e., we first replace h by $h + k$, $h - k$, getting

$$f(c + h + k) = f(c - h - k) + 2(h + k)\varphi(c), \tag{7.3}$$

$$f(c + h - k) = f(c - h + k) + 2(h - k)\varphi(c), \tag{7.4}$$

then form: (7.3) + (7.4) − twice (7.2). This eliminates $\varphi(c)$, giving

$$f(c + h + k) + f(c + h - k) - 2f(c + h)$$
$$= f(c - h + k) + f(c - h - k) - 2f(c - h).$$

Here write $c = h$, then put $2h = x$; we get

$$f(x + k) + f(x - k) - 2f(x) = f(k) + f(-k) - 2f(0). \tag{7.5}$$

Replace x by $x + k$; the right member, being independent of x, remains the same, and we have

$$f(x + 2k) + f(x) - 2f(x + k) = f(k) + f(-k) - 2f(0). \tag{7.6}$$

Subtracting (7.5) from (7.6), we obtain finally:

$$f(x + 2k) - 3f(x + k) + 3f(x) - f(x - k) = 0; \tag{7.7}$$

i.e., *the third difference of $f(x)$ is zero for any arguments in arithmetic progression.*

The only condition for the validity of these calculations is that all the quantities which appear as argument-values of f shall belong to the interval $(-1, 1)$. For this, it will readily be seen to be sufficient that the argument-values in the final result (7.7) belong to this interval.

By successive application of (7.7), we see that the values of $f(x)$ belonging to any arithmetic progression of values of x are all uniquely determined if $f(x)$ is given for any three consecutive values of x in the progression. Now any quadratic polynomial, or parabola,

$$y = ax^2 + bx + c, \tag{7.8}$$

obviously obeys the condition (7.7) of a vanishing third difference, and furthermore can be made to take any given values for three different given values of x by proper (unique) determination of a, b, c. It follows that the points of $y = f(x)$ which belong to any number of values of x in arithmetic progression lie upon a parabola (7.8).[10]

Let $x = p/q$ denote any fixed rational value of x in the interval $(-1, 1)$. The points $x = i/q$ $(i = 0, \pm 1, \cdots, \pm q)$ form an arithmetic progression including p/q. Hence the point $[p/q, f(p/q)]$ lies on a parabola (7.8) with all the points $[i/q, f(i/q)]$. This parabola is uniquely determined by its three points $[0, f(0)]$, $[1, f(1)]$, $[-1, f(-1)]$ $(i = 0, q, -q)$; it is therefore independent of the value of q. In summary, we have

$$f(x) = ax^2 + bx + c \tag{7.9}$$

for all rational values p/q of x.

The postulated continuity of $f(x)$ now implies the quadratic (or linear, or constant) form (7.9) for $f(x)$ for *all* values of x; i.e., $y = f(x)$ is a parabola (or, perhaps, a two-fold straight line).

Brooklyn, N. Y.

[10] Which reduces to a straight line if $a = 0$.

A MODIFIED MOMENT PROBLEM IN TWO VARIABLES

By Stefan Bergman and W. T. Martin

1. In the consideration of the moment problem

$$\int_0^1 \Phi(v) v^{\lambda_n} dv = Y_n \qquad (n = 1, 2, \cdots), \tag{1.1}$$

where $\{\lambda_n\}$ and $\{Y_n\}$ are given sequences of complex numbers with $\mathrm{Re}\,(\lambda_n) > -\frac{1}{2}$ and Φ is sought as a function of $\mathfrak{L}^2$, two cases arise, namely, the homogeneous case and the non-homogeneous case (in the homogeneous case one naturally seeks a function $\Phi \not\equiv 0$). The homogeneous problem is of especial importance since its solution answers the question of the closure of the set $\{v^{\lambda_n}\}$ over $(0, 1)$, the set $\{v^{\lambda_n}\}$ being not closed or closed according as the system (1.1) with $Y_n = 0$ $(n = 1, 2, 3, \cdots)$ has or has not a solution $\Phi \not\equiv 0$, $\Phi \in \mathfrak{L}^2$. Two methods of attack have been used in the homogeneous case.[1] The one method, a real variable method, reduces the problem to an infinite system of linear equations in an infinite number of unknowns and the solution is obtained in terms of infinite determinants, which by use of Cauchy's theorem on determinants and other ingenious considerations yields the result that a necessary and sufficient condition for the closure of the set is the divergence of the series:

$$\sum_{n=1}^{\infty} g(\lambda_n), \tag{1.2}$$

where

$$g(\lambda) = \frac{1 + 2\,\mathrm{Re}\,(\lambda)}{1 + |\lambda|^2}. \tag{1.3}$$

The other method of attack uses Fourier transforms in the complex plane to reduce the problem to a problem on the zeros of analytic functions of a certain class; again the same condition is obtained. Clearly both methods are applicable to the non-homogeneous case.

In this paper we shall be concerned with the moment problem in two variables,[2] namely,

$$\int_0^1 \int_0^1 \Phi(v_1, v_2) v_1^{\lambda_n} v_2^{\mu_n} dv_1\, dv_2 = Y_n \qquad (n = 1, 2, \cdots), \tag{1.4}$$

Received September 12, 1939; in revised form, February 7, 1940.

[1] Cf. Müntz [1], Szász [1] for the first method and Paley-Wiener [2], p. 32, for the second method. (The numbers in brackets refer to the bibliography.) The method of attack of the present paper was suggested by a study of the work of Paley and Wiener. The authors are indebted to Professors Hildebrandt, Tamarkin, and Wiener for helpful advice.

[2] It will be clear that the methods are immediately applicable to the corresponding problem in k variables, rather than in two.

where $\{Y_n\}$ is a given sequence of complex numbers and $\{\lambda_n, \mu_n\}$ is a given sequence of pairs[3] of complex numbers with Re $(\lambda_n) > -\frac{1}{2}$, Re $(\mu_n) > -\frac{1}{2}$, and Φ is sought as a function of $\mathfrak{L}^2$,

$$\int_0^1 \int_0^1 |\Phi(v_1, v_2)|^2 \, dv_1 dv_2 < \infty. \tag{1.5}$$

Like other linear problems the theory of the non-homogeneous case depends essentially upon the question of the existence or non-existence of a solution (not identically zero) of the homogeneous problem

$$\int_0^1 \int_0^1 \Phi(v_1, v_2) v_1^{\lambda_n} v_2^{\mu_n} \, dv_1 dv_2 = 0 \qquad (n = 1, 2, \cdots). \tag{1.6}$$

Accordingly we shall begin the investigation with the homogeneous case. Here the situation is different from that in the one variable case. If we suppose that λ_n, μ_n are not given numbers but merely coördinates of a point in the space of two complex variables, then an equation of the form (1.6) for any fixed Φ represents a certain two-dimensional set $\mathfrak{S}^2$. Therefore, there arises the question of what properties this set has. If we write

$$P(z_1, z_2) = \int_0^1 \int_0^1 \Phi(v_1, v_2) v_1^{z_1} v_2^{z_2} \, dv_1 dv_2, \tag{1.7}$$

then by Schwarz's inequality it is easily seen that $P(z_1, z_2)$ is analytic in E[Re $(z_k) > -\frac{1}{2}$ $(k = 1, 2)$], and therefore by using the theory of a. f. of 2 c. v. (analytic functions of two complex variables) we see that $\mathfrak{S}^2 = \mathrm{E}[P(z_1, z_2) = 0]$ and thus $\mathfrak{S}^2$ must be composed of the analytic manifolds along which P vanishes. If we restrict ourselves to the local behavior of $\mathfrak{S}^2$, i. e., if we associate to any $(\lambda_n, \mu_n) \in \mathfrak{S}^2$ a sufficiently small neighborhood, and if we denote by $\mathfrak{S}_n^2 = \mathfrak{S}_n^2(\lambda_n, \mu_n)$ the part of $\mathfrak{S}^2$ in this neighborhood, then we can distinguish two classes of $\mathfrak{S}_n^2$. One can represent the function $P(z_1, z_2)$ in the neighborhood of the point (λ_n, μ_n) in one and essentially only one way as a product of non-decomposable factors[4]

$$P(z_1, z_2) = p_1^{n_1} \cdots p_r^{n_r} q_1^{m_1} \cdots q_s^{m_s} \Omega(z_1, z_2), \tag{1.8}$$

where the p's are analytic functions vanishing at (λ_n, μ_n) and such that $\partial p_j/\partial z_k$ is non-vanishing at (λ_n, μ_n) for $k = 1$ or $k = 2$, while the q's are analytic non-decomposable analytic functions such that $q_j = \partial q_j/\partial z_1 = \partial q_j/\partial z_2 = 0$ at (λ_n, μ_n), and Ω is a regular function which does not vanish in the neighborhood of (λ_n, μ_n). If

$$r = 1, \qquad n_1 = 1, \qquad s = 0, \tag{1.9}$$

then we shall call the point (λ_n, μ_n) a *simple point* (for Φ). In this case $\mathfrak{S}_n^2$ can be represented either in the form $\lambda = \lambda(\mu)$ or $\mu = \mu(\lambda)$. (As a rule both repre-

[3] The double sequence $\{\lambda_n, \mu_m\}$ in which λ_n, μ_m run independently of one another over all numbers of the infinite sequences $\{\lambda_n\}$ $(n = 1, 2, \cdots)$, $\{\mu_m\}$ $(m = 1, 2, \cdots)$, respectively, is a special case of the sequence $\{\lambda_n, \mu_n\}$ considered by us.

[4] Cf. Osgood [1], Chapter 2.

sentations are possible.) If (1.9) does not occur, then we shall call the point (λ_n, μ_n) a *multiple point*; again the behavior of $\mathfrak{S}_n^2$ can be described. In this case either there are segments of several analytic surfaces through (λ_n, μ_n) if $n_j = 1, r > 1, s = 0$, or there are several such analytic surfaces counted multiply if $n_j > 1$, or the surfaces have branch-points if $s > 0, r = 0$, or finally different combinations of these may occur. In order to save space we shall not discuss the geometric structure of these possible cases; rather, we shall characterize a multiple point briefly by the associated development (1.8). We might point out, however, that for the homogeneous moment problem (1.6) a simple point (λ_n, μ_n) for a solution Φ means that either

$$\int_0^1 \int_0^1 \Phi(v_1, v_2) v_1^{\lambda_n} v_2^{\mu_n} \log v_1 \, dv_1 \, dv_2 \neq 0$$

or

$$\int_0^1 \int_0^1 \Phi(v_1, v_2) v_1^{\lambda_n} v_2^{\mu_n} \log v_2 \, dv_1 \, dv_2 \neq 0,$$

while in the case of a multiple point not only both these but also possibly other expressions of the form

$$\int_0^1 \int_0^1 \Phi(v_1, v_2) v_1^{\lambda_n} v_2^{\mu_n} (\log v_1)^j (\log v_2)^k \, dv_1 \, dv_2 \equiv A_{jk}$$

vanish. We shall call the *order* of the point (λ_n, μ_n) the largest number $\nu - 1$ for which all the expressions A_{jk} vanish, where (j, k) range over the values $(\nu, 0), (\nu - 1, 1), \cdots, (0, \nu)$.

If one not only asks whether a given system of the form (1.6) has a solution but also draws into consideration the segments $\mathfrak{S}_n^2$ for a given solution (assuming that a solution exists), then one can distinguish two categories of points:[5]

(1) Either the point (λ_ν, μ_ν) possesses the property that besides the solution Φ of the problem there exists a function Φ' which satisfies the equations (1.6) for $n \neq \nu$ with the *same* $\mathfrak{S}_n^2$ while the ν-th equation in (1.6) either is not satisfied for Φ' if (λ_n, μ_n) is a simple point for Φ, or if (λ_n, μ_n) is a multiple point for Φ, then at least one of the numbers n_j or m_k corresponding to (1.8) must be diminished by at least one.

(2) Or such a solution Φ' does not exist, that is, every function Φ^* satisfying the equations (1.6) for all n except a finite number of n, necessarily satisfies (1.6) also in a group of points $(\lambda_{\nu_k}, \mu_{\nu_k})$ $(k = 1, 2, \cdots, p)$ which includes (λ_ν, μ_ν). (Obviously in the case of a multiple point (λ_n, μ_n) we may replace non-fulfillment of (1.6) for the point by the lowering of the degree of the point.) Such a group of points $\{\lambda_{\nu_k}, \mu_{\nu_k}\}$ is termed *related* for our solution Φ.

The association with every solution Φ of a function P by means of the correspondence

$$\Phi(v_1, v_2) \to P(z_1, z_2) = \int_0^1 \int_0^1 \Phi(v_1, v_2) v_1^{z_1} v_2^{z_2} \, dv_1 \, dv_2$$

[5] Cf. §2, especially Definition 2.1.

has a serious defect. We cannot conversely immediately associate with every function P (of a suitable class) a solution Φ of the moment problem, and therefore we cannot draw conclusions in the reverse direction $P \to \Phi$, i.e., it is not possible to deduce from the existence of a function $P(z_1, z_2)$ of a suitable class and vanishing in $\{\lambda_n, \mu_n\}$ $(n = 1, 2, \cdots)$ the existence of a solution Φ of (1.6). The second method of attack mentioned in connection with the one variable problem removes this defect. That is, by use of Fourier transforms it is possible to obtain a complete equivalence. We shall now proceed to obtain this equivalence.

Before taking Fourier transforms it is convenient to map the square $0 < v_k < 1$ into the quadrant $q^2 = E[-\infty < u_k < 0, k = 1, 2]$. For this purpose let us write $v_k = e^{u_k}$. Then (1.4) and (1.5) become

$$\int_{-\infty}^{0}\int_{-\infty}^{0} \Phi(e^{u_1}, e^{u_2})e^{u_1(\lambda_n+1)+u_2(\mu_n+1)}\,du_1\,du_2 = Y_n \qquad (n = 1, 2, \cdots),$$

$$\int_{-\infty}^{0}\int_{-\infty}^{0} |\Phi(e^{u_1}, e^{u_2})|^2 e^{u_1+u_2}\,du_1\,du_2 < \infty.$$

Upon writing

$$\varphi(u_1, u_2) = 2\pi\Phi(e^{u_1}, e^{u_2})e^{\frac{1}{2}(u_1+u_2)}, \tag{1.10}$$

we see at once the truth of the following lemma.

LEMMA 1. *A necessary and sufficient condition for the existence of a function Φ satisfying* (1.4) *and* (1.5) *is that a function φ exists satisfying*

$$\int_{-\infty}^{0}\int_{-\infty}^{0} \varphi(u_1, u_2)e^{(\lambda_n+\frac{1}{2})u_1+(\mu_n+\frac{1}{2})u_2}\,du_1\,du_2 = 2\pi Y_n \qquad (n = 1, 2, \cdots), \tag{1.11}$$

$$\int_{-\infty}^{0}\int_{-\infty}^{0} |\varphi(u_1, u_2)|^2\,du_1\,du_2 < \infty. \tag{1.12}$$

The functions φ and Φ are related as in (1.10).

In §3 we prove the following result on Fourier transforms in the complex domain.[6]

THEOREM 1. *The two following classes of analytic functions are identical:*
the class $\mathfrak{W}_1$ of all functions $f(z_1, z_2)$ analytic in the quarter-space $\mathfrak{Q}^4 = E[\mathrm{Im}\,(z_k) > 0\ (k = 1, 2)]$ and such that

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |f(x_1 + iy_1, x_2 + iy_2)|^2\,dx_1\,dx_2 < c = c_f < \infty \qquad (0 < y_k); \tag{1.13}$$

the class $\mathfrak{W}_2$ of all functions defined by

$$f(z_1, z_2) = \frac{1}{2\pi}\int_{-\infty}^{0}\int_{-\infty}^{0} \varphi(u_1, u_2)e^{-i(u_1z_1+u_2z_2)}\,du_1\,du_2, \qquad (z_1, z_2) \in \mathfrak{Q}^4, \tag{1.14}$$

where $\varphi(u_1, u_2)$ belongs to $\mathfrak{L}^2$ over the quadrant $q^2 = E[-\infty < u_k < 0, k = 1, 2]$.

[6] This theorem generalizes to two variables (and its truth was suggested by) results obtained by Bochner [1] (see especially p. 660) and Paley and Wiener [2] (see pp. 8–9).

We shall have

$$f_+(x_1, x_2) = \underset{y_1, y_2 \to 0^+}{\text{l.i.m.}} f(x_1 + iy_1, x_2 + iy_2)$$

(1.15)

$$= \underset{A \to \infty}{\text{l.i.m.}} \frac{1}{2\pi} \int_{-A}^{0} \int_{-A}^{0} \varphi(u_1, u_2) e^{-i(u_1 x_1 + u_2 x_2)} \, du_1 \, du_2 .$$

On the basis of this theorem and Lemma 1 we obtain at once the following result.

LEMMA 2. *A necessary and sufficient condition for the existence of a function* φ *satisfying* (1.11) *and* (1.12) *is that there exist a function* $f \in \mathfrak{W}_1$ *and satisfying*

(1.16) $$f[(\tfrac{1}{2} + \lambda_n)i, (\tfrac{1}{2} + \mu_n)i] = Y_n \qquad (n = 1, 2, \cdots).$$

The functions are related as in (1.14).

Thus the moment problem in (1.4) and (1.5) is completely equivalent to an interpolation problem of the form (1.16) in the class of analytic functions belonging to $\mathfrak{W}_1$.

2. We shall consider in this paper only such solutions Φ for which the points $\{\lambda_n, \mu_n\}$ are unrelated in the sense indicated in §1. In order to give necessary and sufficient conditions for the existence of such solutions it is necessary to consider the manifold $\mathfrak{S}^2$ along which (λ, μ) can vary, in the large and not just in the neighborhood of a given point (λ_n, μ_n) (i.e., not just $\mathfrak{S}_n^2$). (When we say "in the large" we mean throughout the region E[Re $(\lambda) > -\frac{1}{2}$, Re $(\mu) > -\frac{1}{2}$].) If we interpret from the function-theoretic point of view, the fact that the point (λ_ν, μ_ν) is unrelated to the other points (λ_n, μ_n), $n \neq \nu$, means that the associated function P has a factor P_ν which vanishes at (λ_ν, μ_ν), but does not vanish at (λ_n, μ_n), $n \neq \nu$. In the case of a group $\mathfrak{G}^0$ of points $(\lambda_{\nu_k}, \mu_{\nu_k})$ $(k = 1, 2, \cdots, p)$ which are related there does not exist a factor which vanishes in some (but not all) of the points of $\mathfrak{G}^0$, but there exists a factor which vanishes in all points of $\mathfrak{G}^0$, and no point (λ_n, μ_n), $(\lambda_n, \mu_n) \notin \mathfrak{G}^0$. In order to formulate this situation precisely we shall first give an exact definition of unrelated function-zeros (a slightly different definition was previously given by Bergman, [1], [2], p. 146).

DEFINITION 2.1. Let $\mathfrak{R}^4$ be a four-dimensional region and let $\mathfrak{K}(\mathfrak{R}^4)$ be a class of a. f. of 2 c. v. in $\mathfrak{R}^4$. We say that a function $f \in \mathfrak{K}(\mathfrak{R}^4)$ possesses a set $\{\alpha^n\} = \{\alpha_1^n, \alpha_2^n\}$ $(n = 1, 2, \cdots)$, $(\alpha^n) \in \mathfrak{R}^4$, of unrelated zero points in the $\mathfrak{K}(\mathfrak{R}^4)$ sense if for every point (α^n) there exists a function $f_n \in \mathfrak{K}(\mathfrak{R}^4)$ and vanishing at (α^n) such that

(1) $f/\prod_{n=1}^{N} f_n \in \mathfrak{K}(\mathfrak{R}^4)$ for every positive integer N,

(2) $f(\alpha_1^n, \alpha_2^n)/\prod f_{n_\nu}(\alpha_1^n, \alpha_2^n) \neq 0$, where the product is taken over all the func-

tions f_{n_r} of the set $\{f_n\}$ which vanish at (α^n). (Clearly there can be only a finite number of such f_{n_r}.)[7]

(3) $f/f_n = 0$ at the points $\{\alpha^\nu\}$, $\nu \neq n$.

(4) f has no zeros other than points at which some of the f_n vanish.

Remark. It can occur that several f_n are identical. In such a case in all formulas and considerations we assume that the functions f_n occur multiply a number of times corresponding to their multiplicities.

Leaving for consideration at another time the investigation of the case of solutions Φ with related points, we shall give criteria for the existence of solutions with unrelated points. By using the equivalence between the existence of a solution of the moment problem (1.6) and the existence of the associated function f of 2 c. v. possessing zeros in the points $\{(\lambda_n + \frac{1}{2})i, (\mu_n + \frac{1}{2})i\}$ (see Lemmas 1 and 2), we shall show that our question is equivalent to the following problem.

PROBLEM 2.1. *Let us consider a set* $\{\alpha^n\} = \{\alpha_1^n, \alpha_2^n\}$ $(n = 1, 2, \cdots)$ *of points in the unit bicylinder* $\mathfrak{B}^4 = \mathrm{E}[| z_k | < 1 \; (k = 1, 2)]$ *and let* $\mathfrak{L}^2(\mathfrak{B}^4)$ *be the class of functions* $F(z_1, z_2)$ *analytic in* $\mathfrak{B}^4$ *and such that*

$$\int_0^{2\pi}\int_0^{2\pi} | F(r_1 e^{i\theta_1}, r_2 e^{i\theta_2}) |^2 d\theta_1 \, d\theta_2 < c = c_F < \infty \quad (0 < r_k < 1). \tag{2.1}$$

We seek to determine necessary and sufficient conditions for the existence of a function $F \in \mathfrak{L}^2(\mathfrak{B}^4)$ *having the points* $\{\alpha^n\}$ $(n = 1, 2, \cdots)$ *as unrelated zeros in the* $\mathfrak{L}^2(\mathfrak{B}^4)$ *sense.*

The equivalence of our question and Problem 2.1 follows immediately in view of Lemmas 1 and 2 from the following theorem which we shall prove in §4.

THEOREM 2. *The relation*

$$F(\zeta_1, \zeta_2) = -f\left(i\,\frac{1 - \zeta_1}{1 + \zeta_1}, i\,\frac{1 - \zeta_2}{1 + \zeta_2}\right)\frac{4}{(1 + \zeta_1)(1 + \zeta_2)} \tag{2.2}$$

forms a one-to-one reversible correspondence between a function $f(z_1, z_2) \in \mathfrak{W}_1$ *having unrelated zeros at points* $\{t^n\}$ $(n = 1, 2, \cdots)$ *and a function* $F(\zeta_1, \zeta_2) \in \mathfrak{L}^2(\mathfrak{B}^4)$ *having unrelated zeros at points*

$$\{\alpha_1^n, \alpha_2^n\} = \left\{\frac{1 + it_1^n}{1 - it_1^n}, \frac{1 + it_2^n}{1 - it_2^n}\right\}. \tag{2.3}$$

Before giving criteria for Problem 2.1 let us make a remark concerning the form of the criteria. In the case of one variable it is possible to give necessary and sufficient conditions for the solution of the homogeneous moment problem either in the form

$$\lim_{n\to\infty} \gamma_n(\lambda_1, \cdots, \lambda_n) < \infty \tag{2.4}$$

[7] Otherwise there would be an infinite number of zero-surfaces having a limit point α in the interior of the domain, and this point α would be an essential singularity.

or in the form

$$\sum_{n=1}^{\infty} g(\lambda_n) < \infty, \tag{2.5}$$

where $g(\lambda)$ depends only upon λ. The essential content of the papers of Müntz and Szász consists in reducing the condition (2.4) to a condition of the form (2.5). In the case of two variables the classical method of the theory of infinite systems of linear equations in an infinite number of (real) variables or the theory of interpolation of a. f. of 2 c. v. gives us a criterion of the form (2.4). Our task will be to give a criterion of the form (2.5) for the case of unrelated zeros. We obtain criteria for Problem 2.1 from the following theorem whose proof will be given in §5.

THEOREM 3. *Let $\mathfrak{N}(\mathfrak{B}^4)$ be the class of functions $F(\zeta_1, \zeta_2)$ analytic in $\mathfrak{B}^4$ and such that $|F| \leqq 1$ in $\mathfrak{B}^4$. A sufficient condition for the existence of a function $F \in \mathfrak{N}(\mathfrak{B}^4)$ non-vanishing at the origin and having a set $\{\alpha^n\}$ as unrelated zeros in the $\mathfrak{N}(\mathfrak{B}^4)$-sense is that*

$$\sum_{n=1}^{\infty} a(\alpha^n) < \infty, \tag{2.6}$$

where

$$a(\alpha) = -\log A(\alpha) \tag{2.7}$$

and

$$A(\alpha) = \overline{\mathrm{Bd}}_{h \in \mathfrak{G}_\alpha} |h(0, 0)| \tag{2.8}$$

and $\mathfrak{G}_\alpha$ designates the totality of functions $h \in \mathfrak{N}(\mathfrak{B}^4)$ and vanishing at (α).

The same result formulated in terms of the moment problem becomes

THEOREM 3a. *A sufficient condition for the existence of a solution $\Phi \in \mathfrak{L}^2$ ($\Phi \not\equiv 0$) of the homogeneous moment problem* (1.6) *for which $\{\lambda_n, \mu_n\}$ $(n = 1, 2, \cdots)$ forms a set of unrelated points consists in the convergence of the left member of* (2.6) *with*

$$\{\alpha_1^n, \alpha_2^n\} = \left\{\frac{1-2\lambda_n}{3+2\lambda_n}, \frac{1-2\mu_n}{3+2\mu_n}\right\}. \tag{2.9}$$

COROLLARY. *A necessary condition for the closure $\mathfrak{L}^2$ of the set $\{v_1^{\lambda_n}, v_2^{\mu_n}\}$ over $0 < v_k < 1$ $(k = 1, 2)$ is the divergence of the left member of* (2.6) *with the $\{\alpha^n\}$ defined as in* (2.9).

Remark. Since the class of functions $F(\zeta_1, \zeta_2)$, where $F/(2\pi) \in \mathfrak{L}^2(\mathfrak{B}^4)$, includes the class $\mathfrak{N}(\mathfrak{B}^4)$, it is easily seen that

$$A(\alpha) \leqq [1 - (1 - |\alpha_1|^2)(1 - |\alpha_2|^2)]^{\frac{1}{2}}. \tag{2.10}$$

In order to give a necessary condition we need certain notions from a previous paper (Bergman [2], pp. 138, 154) which we shall indicate here briefly. In the

case of one variable the statement of the necessary condition is based upon Jensen's formula. In order to obtain a certain generalization of this formula we must introduce (in the case of a bicylinder) *double harmonic functions*, i.e., functions which satisfy the equations

$$\frac{\partial^2 u}{\partial x_k^2} + \frac{\partial^2 u}{\partial y_k^2} = 0 \qquad (k = 1, 2). \tag{2.11}$$

A double harmonic function can be represented in the bicylinder $\mathfrak{B}^4_{r_1 r_2} = \mathrm{E}[\,|z_k| < r_k]$ by the formula

$$u(R_1 e^{i\Phi_1}, R_2 e^{i\Phi_2}) = \int_0^{2\pi}\int_0^{2\pi} u(r_1 e^{i\varphi_1}, r_2 e^{i\varphi_2}) \prod_{k=1}^{2} P(R_k e^{i\Phi_k}, r_k e^{i\varphi_k})\, d\varphi_k, \tag{2.12}$$

where

$$P(Re^{i\Phi}, re^{i\varphi}) = \frac{1}{2\pi}\,\frac{r^2 - R^2}{r^2 + R^2 - 2rR\cos(\varphi - \Phi)}.$$

Along with double harmonic functions there was introduced in the previous paper the *double harmonic Green's function* for an analytic function $h(z_1, z_2)$ regular in $\overline{\mathfrak{B}}^4_{r_1 r_2}$ defined by

$$\Gamma(z_1, z_2; h; r_1, r_2) \equiv \Gamma(z_1, z_2; h; \mathfrak{B}^4_{r_1 r_2}) = -\log|h(z_1, z_2)| + \mathfrak{D}(z_1, z_2), \tag{2.13}$$

where $\mathfrak{D}(z_1, z_2)$ designates the double harmonic function in $\mathfrak{B}^4_{r_1 r_2}$ which assumes the values $\log|h(z_1, z_2)|$ on $\mathrm{E}[|z_k| = r_k]$. We obtain the following necessary condition for Problem 2.1 in terms of these notions; the proof will be given in §6.

THEOREM 4. *Let $\mathrm{I}(\mathfrak{B}^4)$ be the class of functions F analytic in $\mathfrak{B}^4$ and such that*

$$\int_0^{2\pi}\int_0^{2\pi} \log^+|F(r_1 e^{i\theta_1}, r_2 e^{i\theta_2})|\, d\theta_1\, d\theta_2 < c = c_r < \infty \qquad (0 < r_k < 1). \tag{2.14}$$

A necessary condition for the existence of a function $F \in \mathrm{I}(\mathfrak{B}^4)$ non-vanishing at the origin and having a set $\{\alpha^n\}$ as unrelated zeros in the $\mathrm{I}(\mathfrak{B}^4)$-sense is that

$$\sum_{n=1}^{\infty} C(\alpha^n) < \infty, \tag{2.15}$$

where

$$C(\alpha) = \lim_{r\to 1} C(\alpha, r, r) \tag{2.16}$$

and

$$C(\alpha, r, r) = \underset{h \in \mathfrak{H}_\alpha}{\underline{\mathrm{Bd}}}\, \Gamma(0, 0; h; \mathfrak{B}^4_{rr}) \tag{2.17}$$

and $\mathfrak{H}_\alpha$ is the class of functions $h \in \mathrm{I}(\mathfrak{B}^4)$ and vanishing at (α).

Remark. In another paper (Bergman [3]) it is proved that (2.15) is also a sufficient condition for the existence of a double harmonic function possessing

indicated singularities at the points (α^n). Since the class of double harmonic functions is larger than the class of biharmonic functions, there arises the task of finding in addition to (2.15) sufficient conditions to ensure that the double harmonic function so obtained will be biharmonic. In such a case the function $\exp(-b - ic)$, where b is this biharmonic function and c is the conjugate to b, will be the desired a. f. of 2 c. v. Such considerations belong to the theory of a. f. of 2 c. v. and lie beyond the realm of this paper. They will be treated elsewhere.

In the case of one complex variable, where it is not necessary to introduce a larger class (than the class of harmonic functions), the condition (2.15) becomes

$$-\sum_{n=1}^{\infty} \log |\alpha^n| < \infty. \tag{2.18}$$

This is the known necessary and sufficient condition for the existence of a function of the class $\mathfrak{L}^2(\mathfrak{B}^4)$ having zeros at the points (α^n). On the other hand (2.6) also yields in the case of one variable the same condition (2.18) since $A(\alpha) = |\alpha|$. If we substitute $\alpha^n = (1 - 2\lambda_n)/(3 + 2\lambda_n)$ (as in (2.9)) into (2.18), this yields at once the classical result that the necessary and sufficient condition for the closure of the set $\{v^{\lambda_n}\}$ over (0, 1) is the divergence of the series

$$\sum_{n=1}^{\infty} \frac{1 + 2\,\mathrm{Re}\,(\lambda_n)}{1 + |\lambda_n|^2} \tag{2.19}$$

(see e.g. Paley and Wiener [1], pp. 761–764).

In terms of the moment problem Theorem 4 becomes

THEOREM 4a. *A necessary condition for the existence of a solution* $\Phi \in \mathfrak{L}^2(\mathfrak{B}^4)$, $\Phi \not\equiv 0$, *of the homogeneous moment problem* (1.6) *for which* $\{\lambda_n, \mu_n\}$ $(n = 1, 2, \cdots)$ *forms a set of unrelated points consists in the convergence of the left member of* (2.15) *with* $\{\alpha^n\}$ *defined as in* (2.9).

3. In this section we give a proof of Theorem 1 and then derive a corollary which we later use in proving Theorem 2. We base the proof on the following theorem of Bochner [2].

THEOREM A. *Let* $\mathfrak{K}^2$ *be an open point set in the* y_1y_2*-plane and let* $\mathfrak{T}^4 = \mathfrak{T}^4_{\mathfrak{K}}$ *be the region consisting of all points* (z_1, z_2) *for which* $(y_1, y_2) \in \mathfrak{K}^2$ $(-\infty < x_k < \infty;$ $k = 1, 2)$. *A necessary and sufficient condition that a function* $f(z_1, z_2)$ *analytic in* $\mathfrak{T}^4$ *be of integrable square uniformly in* $\mathfrak{T}^4$, *that is, that*

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |f(x_1 + iy_1, x_2 + iy_2)|^2\, dx_1\, dx_2 \leqq c < \infty, \qquad (y) \in \mathfrak{K}^2, \tag{3.1}$$

is that f *have the representation*

$$f(z_1, z_2) \sim \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \psi(u_1, u_2)e^{-i(u_1z_1+u_2z_2)}\, du_1\, du_2, \qquad (z) \in \mathfrak{T}^4, \tag{3.2}$$

where for each $(y) \in \mathfrak{H}^2$ *the function*

$$\psi(u_1, u_2)e^{u_1y_1+u_2y_2} \tag{3.3}$$

belongs to $\mathfrak{L}^2$ *in the* u_1u_2*-space.*

From Bochner's theorem it follows at once that the class $\mathfrak{W}_2 \subset \mathfrak{W}_1$. However, since the direct proof of this part of the theorem is so brief, we shall give it here. If $\varphi(u_1, u_2) \in \mathfrak{L}^2$ over $q^2 = \mathrm{E}[-\infty < u_k < 0, k = 1, 2]$, then the integral on the right side of (1.14) converges absolutely and uniformly in every finite region contained in $\mathfrak{Q}^4$ as one sees by applying the Schwarz inequality:

$$\left[\int_{-\infty}^{0}\int_{-\infty}^{0} | \varphi(u_1, u_2)e^{-i(u_1z_1+u_2z_2)} | du_1 du_2\right]^2$$
$$\leqq \int_{-\infty}^{0}\int_{-\infty}^{0} | \varphi(u_1, u_2) |^2 du_1 du_2 \cdot \int_{-\infty}^{0}\int_{-\infty}^{0} e^{2(u_1y_1+u_2y_2)} du_1 du_2 = \text{const.}\ \frac{1}{4y_1y_2}.$$

Thus the function $f(z_1, z_2)$ defined by (1.14) is analytic in $\mathfrak{Q}^4$. Also by Plancherel's theorem[8]

$$\begin{aligned} &\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} | f(x_1 + iy_1, x_2 + iy_2) |^2 dx_1 dx_2 \\ &= \int_{-\infty}^{0}\int_{-\infty}^{0} | \varphi(u_1, u_2) |^2 e^{2(u_1y_1+u_2y_2)} du_1 du_2 \leqq \int_{-\infty}^{0}\int_{-\infty}^{0} | \varphi(u_1, u_2) |^2 du_1 du_2 \end{aligned} \tag{3.4}$$

for $y_k > 0$ $(k = 1, 2)$. Thus $\mathfrak{W}_2 \subset \mathfrak{W}_1$.

Next we show that $\mathfrak{W}_1 \subset \mathfrak{W}_2$. By Bochner's theorem, if $f \in \mathfrak{W}_1$, then f has the representation (3.2) in $\mathfrak{Q}^4$, where (by Plancherel's theorem)

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} | \psi(u_1, u_2) |^2 e^{2(u_1y_1+u_2y_2)} du_1 du_2 < c \qquad (0 < y_k). \tag{3.5}$$

The boundedness of the integral for $0 < y_k$ implies that $\psi(u_1, u_2) \equiv 0$ for (u_1, u_2) outside the quadrant q^2. For suppose $\psi \not\equiv 0$ in some rectangle lying *outside* q^2. Then

$$\int_a^b\int_\alpha^\beta | \psi(u_1, u_2) |^2 du_1 du_2 = \mathfrak{J} > 0$$

for some a, b, α, β with, say, $0 < a < b$ and $\alpha < \beta$ (α may be positive, zero, or negative). Then

$$\int_a^b\int_\alpha^\beta | \psi(u_1, u_2) |^2 e^{2(u_1y_1+u_2y_2)} du_1 du_2 > \mathfrak{J}e^{2(ay_1+\alpha y_2)}$$

[8] For various theorems and results on Fourier transforms in the real space see, for example, Wiener, *The Fourier Integral*, Cambridge, Eng., 1933. Wiener shows that "no property of the Lebesgue integral not explicitly involving a specified number of dimensions in the enunciation is restricted to integrals in any particular number of dimensions". (See p. 17.)

and as $y_1 \to \infty$ (for fixed positive y_2) this contradicts the boundedness of (3.5). Thus $\psi(u_1, u_2) \equiv 0$ for (u_1, u_2) outside q^2.

Again, using (3.5) we see by monotone convergence as[9] $\{y\} \to 0^+$ that $\psi(u_1, u_2) \in \mathfrak{L}^2$ over q^2. Thus we have proved that $\mathfrak{W}_1 \subset \mathfrak{W}_2$.

Next we show that (1.15) holds. For this purpose let us denote by f_+^* the function

$$f_+^*(x_1, x_2) = \operatorname*{l.i.m.}_{A\to\infty} \frac{1}{2\pi} \int_{-A}^{0}\int_{-A}^{0} \varphi(u_1, u_2) e^{-i(u_1x_1+u_2x_2)}\, du_1\, du_2 \tag{3.6}$$

which clearly exists and belongs to $\mathfrak{L}^2$ since $\varphi \in \mathfrak{L}^2$. Using Plancherel's theorem, we see that

$$\begin{aligned}
&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |f(x_1 + iy_1, x_2 + iy_2) - f_+^*(x_1, x_2)|^2\, dx_1\, dx_2 \\
&= \int_{-\infty}^{0}\int_{-\infty}^{0} |\varphi(u_1, u_2)|^2 \cdot |e^{u_1y_1+u_2y_2} - 1|^2\, du_1\, du_2 \\
&= \left[\int_{-\infty}^{-A}\int_{-\infty}^{0} + \int_{-A}^{0}\int_{-\infty}^{-A} + \int_{-A}^{0}\int_{-A}^{0}\right][|\varphi|^2 \cdot |e^{u_1y_1+u_2y_2} - 1|^2\, du_1\, du_2] \qquad (3.7)\\
&\leqq 4\left[\int_{-\infty}^{-A}\int_{-\infty}^{0} + \int_{-A}^{0}\int_{-\infty}^{-A}\right][|\varphi|^2\, du_1\, du_2] \\
&\qquad + \left[\int_{-A}^{0}\int_{-A}^{0} |\varphi|^2\, du_1\, du_2\right][\max_{\substack{-A\leqq u_1\leqq 0\\ -A\leqq u_2\leqq 0}} |e^{u_1y_1+u_2y_2} - 1|^2].
\end{aligned}$$

If $\epsilon > 0$ there exist an $A_0 = A_0(\epsilon)$ and a $y_0 = y_0(A_0, \epsilon)$ such that

$$\begin{aligned}
&\int_{-\infty}^{-A_0}\int_{-\infty}^{0} |\varphi|^2\, du_1\, du_2 < \epsilon, \qquad \int_{-A_0}^{0}\int_{-\infty}^{-A_0} |\varphi|^2\, du_1\, du_2 < \epsilon, \\
&\max_{\substack{-A_0\leqq u_1\leqq 0\\ -A_0\leqq u_2\leqq 0}} |e^{u_1y_1+u_2y_2} - 1|^2 < \frac{\epsilon}{\int_{-\infty}^{0}\int_{-\infty}^{0} |\varphi|^2\, du_1\, du_2} \quad \text{for} \quad 0 < y_k < y_0.
\end{aligned} \tag{3.8}$$

Thus

$$\operatorname*{l.i.m.}_{(y)\to 0^+} f(x_1 + iy_1, x_2 + iy_2) = f_+^*(x_1, x_2) \tag{3.9}$$

and (1.15) holds.

This concludes the proof of Theorem 1.

We also need the following form of Cauchy's theorem

[9] We shall write $(y) \to 0^+$ to mean that y_1, y_2 run over the values $\epsilon_1^{(n)}$, $\epsilon_2^{(n)}$ respectively, where $\epsilon_k^{(n)} > \epsilon_k^{(n+1)} > 0$ and $\epsilon_k^{(n)} \to 0$. When we write $\lim_{(y)\to 0^+}$, we mean that the limit exists and is the same for all such approaches.

COROLLARY 1. *If $f(z_1, z_2) \in \mathfrak{W}_1$, then it has the following representation*

$$f(z_1, z_2) = \left(\frac{1}{2\pi i}\right)^2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{f_+(u_1, u_2)}{(u_1 - z_1)(u_2 - z_2)}\, du_1\, du_2, \qquad (z) \in \mathfrak{Q}^4, \tag{3.10}$$

where f_+ is the function defined by (1.15).

For the proof of the corollary we apply Parseval's theorem[10] to the equation (1.14). Now the function $f_+(t_1, t_2)$ has the Fourier transform $\varphi(u_1, u_2)$ and the function

$$\frac{1}{(2\pi i)^2}\frac{1}{t_1 + iy_1}\frac{1}{t_2 + iy_2}, \tag{3.11}$$

for fixed positive y_1, y_2, has the Fourier transform[11]

$$\begin{cases} \dfrac{1}{2\pi} e^{u_1y_1+u_2y_2} & \text{for } (u_1, u_2) \in q^2, \\ 0 & \text{for } (u_1, u_2) \text{ outside } q^2. \end{cases} \tag{3.12}$$

Thus, applying Parseval's theorem, we see that

$$\begin{aligned} &\frac{1}{2\pi}\int_{-\infty}^{0}\int_{-\infty}^{0} \varphi(u_1, u_2) e^{u_1y_1+u_2y_2} e^{-i(u_1x_1+u_2x_2)}\, du_1\, du_2 \\ &\qquad = \left(\frac{1}{2\pi i}\right)^2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{f_+(t_1, t_2)}{[(x_1 - t_1) + iy_1][(x_2 - t_2) + iy_2]}\, dt_1\, dt_2. \end{aligned} \tag{3.13}$$

On comparing (1.14) and (3.13), we see that f has the representation (3.10) in $\mathfrak{Q}^4$.

4. In this section we give a proof of Theorem 2. First let $f \in \mathfrak{W}_1$. Then f has the representation (3.10). If we write

$$z_k = i\frac{1 - \zeta_k}{1 + \zeta_k}, \qquad u_k = \tan \tfrac{1}{2}\theta_k = i\frac{1 - e^{i\theta_k}}{1 + e^{i\theta_k}} \qquad (k = 1, 2), \tag{4.1}$$

then

$$\frac{1}{u_k - z_k} = i\frac{(1 + \zeta_k)(1 + e^{i\theta_k})}{2(e^{i\theta_k} - \zeta_k)} = i\frac{(1 + \zeta_k)(1 + e^{i\theta_k})}{2e^{i\theta_k}} \sum_{m=0}^{\infty} \zeta_k^m e^{-im\theta_k}, \qquad |\zeta_k| < 1. \tag{4.2}$$

[10] Cf. Paley-Wiener [2], (1.09), p. 2.

[11] See Paley-Wiener [2], p. 5, or note that

$$\begin{aligned} \frac{1}{2\pi}\int_{-\infty}^{0}\int_{-\infty}^{0} \frac{1}{2\pi} e^{u_1y_1+u_2y_2-iu_1t_1-iu_2t_2}\, du_1\, du_2 &= \left(\frac{1}{2\pi}\right)^2 \frac{1}{(y_1 - it_1)(y_2 - it_2)} \\ &= \left(\frac{1}{2\pi i}\right)^2 \frac{1}{(t_1 + iy_1)(t_2 + iy_2)}. \end{aligned}$$

If we map back by means of

$$\zeta_k = \frac{1 + iz_k}{1 - iz_k}, \tag{4.3}$$

then we have $1 + \zeta_k = 2/(1 - iz_k)$ and (4.3) becomes

$$\frac{z_k + i}{u_k - z_k} = i\,\frac{1 + e^{i\theta_k}}{e^{i\theta_k}} \sum_{m=1}^{\infty} \zeta_k^m e^{-im\theta_k} = \frac{2}{1 + i\tan\theta_k} \sum_{m=1}^{\infty} \zeta_k^m e^{-im\theta_k}, \qquad |\zeta_k| < 1. \tag{4.4}$$

Thus by (3.10) and (2.2)

$$\begin{aligned} &f(z_1, z_2)(z_1 + i)(z_2 + i) \\ &\quad = \left(\frac{1}{2\pi i}\right)^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{f_+(u_1, u_2)}{(u_1 - z_1)(u_2 - z_2)} (z_1 + i)(z_2 + i)\, du_1\, du_2 \\ &\quad = F(\zeta_1, \zeta_2) = -\frac{1}{4\pi^2} \int_0^{2\pi} \int_0^{2\pi} f_+(\tan \tfrac{1}{2}\theta_1, \tan \tfrac{1}{2}\theta_2) \\ &\qquad\qquad \cdot (1 - i\tan\theta_1)(1 - i\tan\theta_2) \sum_{m,n=0}^{\infty} \zeta_1^m \zeta_2^n e^{-i(m\theta_1 + n\theta_2)}\, d\theta_1\, d\theta_2 . \end{aligned} \tag{4.5}$$

Since

$$\begin{aligned} &\frac{1}{4} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} |f_+(\tan \tfrac{1}{2}\theta_1, \tan \tfrac{1}{2}\theta_2)|^2 \,|1 - i\tan\tfrac{1}{2}\theta_1|^2 \,|1 - i\tan\tfrac{1}{2}\theta_2|^2\, d\theta_1\, d\theta_2 \\ &\qquad = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |f_+(u_1, u_2)|^2\, du_1\, du_2 < \infty, \end{aligned} \tag{4.6}$$

it follows at once that we may integrate (4.5) term-wise for $(\zeta) \in \mathfrak{B}^4$ and thus

$$F(\zeta_1, \zeta_2) = \sum_{m,n=0}^{\infty} a_{mn} \zeta_1^m \zeta_2^n, \tag{4.7}$$

where

$$\begin{aligned} a_{mn} = -\frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} & f_+(\tan \tfrac{1}{2}\theta_1, \tan \tfrac{1}{2}\theta_2) \\ &\cdot (1 - i\tan\tfrac{1}{2}\theta_1)(1 - i\tan\tfrac{1}{2}\theta_2) e^{-i(m\theta_1 + n\theta_2)}\, d\theta_1\, d\theta_2 , \end{aligned}$$

and since the set $\{(2\pi)^{-1} e^{i(m_1\theta_1 + m_2\theta_2)}\}$ is orthonormal, we have by Bessel's inequality

$$\sum_{m,n=0}^{\infty} |a_{mn}|^2 < \infty.$$

Thus $F(\zeta_1, \zeta_2) \in \mathfrak{L}^2(\mathfrak{B}^4)$.

Conversely if $F(\zeta_1, \zeta_2) \in \mathfrak{L}^2(\mathfrak{B}^4)$, then the function $f(z_1, z_2)$ defined in (2.2) belongs to $\mathfrak{W}_1$. This may be shown by an argument entirely analogous to an

argument given elsewhere in another connection.[12] We shall not give the details here but we shall simply note that if $F(\zeta_1, \zeta_2) \in \mathfrak{N}(\mathfrak{B}^4)$, i.e., if F is analytic in $\mathfrak{B}^4$ and $|F| \leqq 1$ in $\mathfrak{B}^4$, then the function $f(z_1, z_2)$ related to F as in (2.2) is analytic in $\mathfrak{Q}^4$ and $|f(z_1, z_2)| \leqq (|z_1 + i| \cdot |z_2 + i|)^{-1}$ in $\mathfrak{Q}^4$. Thus $f \in \mathfrak{W}_1$ in this case. Let us remark that the only place in which this portion of Theorem 2 was used was in passing from Theorem 3 to Theorem 3a, and for this passage it is only necessary to know that $f \in \mathfrak{W}_1$ whenever $F \in \mathfrak{N}(\mathfrak{B}^4)$.

The fact that F and f have zeros of the desired nature at the points indicated in Theorem 2 follows at once by the mappings (4.1) and (4.3). We shall omit the details.

5. In this section we give a proof of Theorem 3. In order to prove Theorem 3 we must make some remarks concerning $a(\alpha)$ and certain functions $g_\alpha(z_1, z_2)$ which we shall need in our later considerations.

1°. Since $\mathfrak{G}_\alpha$ is a normal family, there exists at least one function $g_\alpha(z_1, z_2)$ of $\mathfrak{G}_\alpha$ such that

$$|g_\alpha(0, 0)| = A(\alpha). \tag{5.1}$$

2°. $g_\alpha(z_1, z_2)$ cannot vanish in the bicylinder $\mathrm{E}[|z_k| < |\alpha_k|]$. For the proof, suppose $g_\alpha(\beta_1, \beta_2) = 0$ where $|\beta_k| < |\alpha_k|$. Then the function $g_\alpha(\beta_1 z_1/\alpha_1, \beta_2 z_2/\alpha_2)$ will belong to $\mathfrak{G}_\alpha$ and will assume in $|z_k| \leqq 1$ values which $g_\alpha(z_1, z_2)$ assumes in $|z_k| \leqq |\beta_k/\alpha_k|$. Since $|\beta_k/\alpha_k| < 1$, it follows that

$$\left| g_\alpha\left(\frac{\beta_1 z_1}{\alpha_1}, \frac{\beta_2 z_2}{\alpha_2}\right)\right| < 1, \qquad |z_k| \leqq 1, \tag{5.2}$$

and thus there exists a constant $\epsilon > 0$ such that

$$(1 + \epsilon) g_\alpha\left(\frac{\beta_1 z_1}{\alpha_1}, \frac{\beta_2 z_2}{\alpha_2}\right) \in \mathfrak{N}(\mathfrak{B}^4)$$

and this contradicts (5.1) since

$$(1 + \epsilon)\, |g_\alpha(0, 0)| > |g_\alpha(0, 0)|. \tag{5.3}$$

3°. If $|\beta_k| < |\alpha_k|$ $(k = 1, 2)$, then

$$A(\beta) \leqq A(\alpha). \tag{5.4}$$

The proof of this is analogous to that of 2°.

4°. If the set $\{\alpha^n\}$ $(n = 1, 2, \cdots)$ has the property that

$$\lim_{n\to\infty} |\alpha_k^n| = 1 \qquad (k = 1, 2), \tag{5.5}$$

[12] The essential difficulty is to show that a certain series can be integrated term-wise and this can be shown by using a method indicated in the paper of Bergman, Bull. de l'Institut Math. de Tomsk, vol. 1(1937), pp. 246 and 247.

then

$$\lim_{n\to\infty} A(\alpha^n) = 1. \tag{5.6}$$

It is obviously necessary only to show that there exists a set of functions $h_n(z_1, z_2)$ of $\mathfrak{N}(\mathfrak{B}^4)$ with $h_n(\alpha_1^n, \alpha_2^n) = 0$ and such that $\lim_{n\to\infty} | h_n(0, 0) | = 1$. The set

$$h_n(z_1, z_2) = \prod_{k=1}^{2} \frac{z_k - \alpha_k^n}{1 - \bar{\alpha}_k^n z_k}$$

forms such a set.

5°. Let

$$\lim_{n\to\infty} A(\alpha^n) = 1. \tag{5.7}$$

Then for every (r_1, r_2), $r_k < 1$, there exists an N such that the functions $g_{\alpha^n}(z_1, z_2)$, $n \geqq N$, do not vanish in $| z_k | \leqq r_k$.

For the proof suppose that the assertion is false. Then there exists an infinite subset $g_{\alpha^m}(z_1, z_2) = c_\mu(z_1, z_2)$, where $m = n_\mu \to \infty$ as $\mu \to \infty$, such that each c_μ vanishes at some point (γ^μ), $| \gamma_k^\mu | \leqq r_k < 1$. Therefore by (5.1), (2.8) and 3°, it follows that

$$A(\alpha^{n_\mu}) = | c_\mu(0, 0) | \leqq A(\gamma^\mu) \leqq A(r) < 1,$$

and this contradicts (5.7).

After these preliminaries we pass to the proof itself. Let us denote by $g_n(z_1, z_2)$ the function $g_{\alpha^n}(z_1, z_2)$ defined in 1° corresponding to (α^n). We shall show that *the function*

$$g(z_1, z_2) = \prod_{n=1}^{\infty} g_n(z_1, z_2) \tag{5.8}$$

belongs to $\mathfrak{N}(\mathfrak{B}^4)$ *and that it possesses the set* $\{\alpha^n\}$ *as a set of unrelated zeros in the* $\mathfrak{N}(\mathfrak{B}^4)$*-sense* (cf. Bergman [2], p. 146). It follows from (2.6) and (2.7) that (5.6) holds and therefore by 5° there exists for every (r_1, r_2), $r_k < 1$, an M_0 such that the functions $g_n(z_1, z_2)$, $n \geqq M_0$, do not vanish in $| z_k | \leqq r_k$. For every M, $M > M_0$, $\prod_{v=M_0}^{M} g_v(z_1, z_2)$ will be a function which does not vanish in $| z_k | \leqq r_k$ and therefore $-\sum_{v=M_0}^{M} \log | g_v(z_1, z_2) |$ will be a positive regular biharmonic function in $| z_k | \leqq r_k$. By (2.12)

$$\begin{aligned} -\sum_{v=M_0}^{M} &\log | g_v(z_1, z_2) | \\ &= -\int_0^{2\pi}\int_0^{2\pi} \sum_{v=M_0}^{M} \log | g_v(r_1 e^{i\varphi_1}, r_2 e^{i\varphi_2}) | \prod_{k=1}^{2} P(z_k, r_k e^{i\varphi_k})\, d\varphi_k \end{aligned} \tag{5.9}$$

for $|z_k| < r_k$. Since the right side is non-negative in $|z_k| \leqq r_k$, it follows from (5.9) and the mean-value property of biharmonic functions that

$$(5.10) \qquad -\sum_{\nu=M_0}^{M} \log |g_\nu(z_1, z_2)| \leqq \left(\prod_{k=1}^{2} \frac{r_k + |z_k|}{r_k - |z_k|}\right)\left(-\sum_{\nu=M_0}^{M} \log A(\alpha^\nu)\right).$$

Thus by our hypothesis (see (2.6) and (2.7)) the series

$$(5.11) \qquad -\sum_{\nu=M_0}^{\infty} \log |g_\nu(z_1, z_2)|$$

converges uniformly in $|z_k| < r_k$ and hence $\prod_{\nu=M_0}^{\infty} g_\nu(z_1, z_2)$ ($\not\equiv 0$) and also (5.8) are analytic in $|z_k| < r_k$. Thus (5.8) is analytic in $\mathfrak{B}^4$. Clearly this function $g(z_1, z_2)$ has unrelated zeros at the points $\{\alpha^n\}$. This concludes the proof of Theorem 3.

6. In this section we give a proof of Theorem 4. In §2 we mentioned double harmonic functions and we noticed that every double harmonic function regular in $|z_k| < r_k$ which assumes values $u(r_1e^{i\varphi_1}, r_2e^{i\varphi_2})$ on the two-dimensional surface $|z_k| = r_k$ can be represented in the form (2.12). With the aid of double harmonic functions we associated with every analytic function $h(z_1, z_2)$ regular in $\overline{\mathfrak{B}}^4_{r_1r_2} = \mathrm{E}[|z_k| \leqq r_k]$ the double harmonic Green's function $\Gamma(z_1, z_2; h; \mathfrak{B}^4_{r_1r_2})$ given by (2.13). We wish to show that $\Gamma(z_1, z_2; h; \mathfrak{B}^4_{r_1r_2})$ possesses the following properties which we need for our proof:

1°. $\Gamma(z_1, z_2; h; \mathfrak{B}^4_{r_1r_2}) \geqq 0$ for $|z_k| < r_k$.

Proof. Let us denote by $\mathfrak{U}^4_A$ the domain $\mathrm{E}[|z_k| < r_k, \log|h(z_1, z_2)| > -A]$ and let us place

$$(6.1) \qquad \mathfrak{D}_A(z_1, z_2) = \int_0^{2\pi}\int_0^{2\pi} [\log|h(r_1e^{i\varphi_1}, r_2e^{i\varphi_2})|]_A \prod_{k=1}^{2} P(z_k, r_ke^{i\varphi_k})\, d\varphi_k,$$

where

$$[n]_A = \begin{cases} n & \text{if } n \geqq -A, \\ -A & \text{if } n \leqq -A. \end{cases}$$

We have then

$$(6.2) \qquad \mathfrak{D}_A(z_1, z_2) \geqq -A, \qquad |z_k| < r_k.$$

Next we show that

$$(6.3) \qquad -\log|h(z_1, z_2)| + \mathfrak{D}_A(z_1, z_2)$$

is non-negative in $\mathfrak{U}^4_A$. In the part of the boundary of $\mathfrak{U}^4_A$,

$$\mathrm{E}[|z_k| \leqq r_k, -\log|h(z_1, z_2)| = A],$$

(6.3) is simply $-A + \mathfrak{D}_A(z_1, z_2)$ which is non-negative by (6.2). The supplementary part of the boundary of $\mathfrak{U}^4_A$ consists of

(6.4) $$E[|z_1| = r_1, |z_2| \leqq r_2, \log|h(z_1, z_2)| > -A]$$

and

(6.5) $$E[|z_1| \leqq r_1, |z_2| = 1, \log|h(z_1, z_2)| > -A].$$

In every point of the two-dimensional surface

(6.6) $$E[|z_k| = r_k, \log|h(z_1, z_2)| > -A]$$

(6.3) is zero. To see that (6.3) is non-negative in (6.4) we note that in every lamina

(6.6) $$E[z_1 = r_1e^{i\varphi_1}, |z_2| \leqq r_2, \log|h(z_1, z_2)| > -A] \ (\varphi_1 \text{ fixed})$$

(6.3) is a harmonic function of x_2, y_2, and since it is non-negative on the boundary (the boundary being part of (6.6)), it is non-negative in (6.6). Since this is true for any real φ_1, we have (6.3) non-negative in (6.4). In an analogous manner we prove that (6.3) is non-negative in (6.5). Thus (6.3) is non-negative on the entire boundary of $\mathfrak{U}_A^4$, and since (6.3) is a double harmonic function, it must be non-negative in $\mathfrak{U}_A^4$, i.e.,

(6.7) $$-\log|h(z_1, z_2)| + \mathfrak{D}_A(z_1, z_2) \geqq 0, \qquad (z) \in \mathfrak{U}_A^4.$$

On the other hand we have

(6.8) $$\Gamma(z_1, z_2; h; \mathfrak{B}_{r_1r_2}^4) = -\log|h(z_1, z_2)| + \lim_{A\to\infty} \mathfrak{D}_A(z_1, z_2).$$

1° follows from (6.7) and (6.8).

2°. $C(\alpha, r_1, r_2) \leqq C(\alpha, R_1, R_2)$ for $r_k < R_k$.

Proof. For every $\epsilon > 0$ there exists a function h_ϵ belonging to $\mathfrak{H}_\alpha$ such that

(6.9) $$C(\alpha, R_1, R_2) = \Gamma(0, 0; h_\epsilon; \mathfrak{B}_{R_1R_2}^4) - \epsilon$$

and therefore

(6.10) $$C(\alpha, R_1, R_2) - C(\alpha, r_1, r_2) \geqq \Gamma(0, 0; h_\epsilon; \mathfrak{B}_{R_1R_2}^4) - \epsilon - \Gamma(0, 0; h_\epsilon; \mathfrak{B}_{r_1r_2}^4).$$

On the other hand

(6.11) $$\Gamma(z_1, z_2; h_\epsilon; \mathfrak{B}_{R_1R_2}^4) - \Gamma(z_1, z_2; h_\epsilon; \mathfrak{B}_{r_1r_2}^4)$$

is a regular double harmonic function in $\mathfrak{B}_{r_1r_2}^4$. Also by 1° we see that

(6.12) $$\Gamma(z_1, z_2; h_\epsilon; \mathfrak{B}_{R_1R_2}^4) \geqq 0, \qquad |z_k| = r_k,$$

and by (2.13)

(6.13) $$\Gamma(z_1, z_2; h_\epsilon; \mathfrak{B}_{r_1r_2}^4) = 0, \qquad |z_k| = r_k,$$

except possibly at the set at which $h_\epsilon(r_1e^{i\varphi_1}, r_2e^{i\varphi_2}) = 0$. Since a regular double harmonic function in a bicylinder $\mathfrak{B}_{r_1r_2}^4$ assumes its minimum on the two-dimensional surface $|z_k| = r_k$, it follows that (6.11) is non-negative in $\mathfrak{B}_{r_1r_2}^4$ and

hence in particular at the point $z_k = 0$. Thus by (6.10) we obtain

$$C(\alpha, R_1, R_2) \geqq C(\alpha, r_1, r_2) - \epsilon \tag{6.14}$$

and since this holds for every positive ϵ, 2° follows.

3°. The limit

$$C(\alpha) = \lim_{r \to 1^-} C(\alpha, r, r) \tag{6.15}$$

exists and is finite.

This follows at once from 2° and the inequality

$$C(\alpha, r_1, r_2) \leqq \log \left| \frac{r_1 r_2}{\alpha_1 \alpha_2} \right|. \tag{6.16}$$

We pass now to the proof of Theorem 4. Let $F \in \mathfrak{l}(\mathfrak{B}^4)$, let $F(0, 0) \neq 0$ and let F have a set $\{\alpha^n\}$ of unrelated zeros. Let $F_1, F_2, \cdots$ be zero-functions of F having the properties $1°, \cdots, 4°$ stated in Definition 2.1. For every $r < 1$ there exists an $N(r)$ such that the functions F_ν, $\nu > N(r)$, do not vanish in $\mathfrak{B}^4_{rr}$. Therefore

$$\frac{F}{\prod_{\nu=1}^{N(r)} F_\nu}$$

is a regular function which does not vanish in $\mathfrak{B}^4_{rr}$. Hence

$$\log | F(z_1, z_2) | - \sum_{\nu=1}^{N(r)} \log | F_\nu(z_1, z_2) | \tag{6.17}$$

and therefore also

$$\log | F(z_1, z_2) | + \sum_{\nu=1}^{N(r)} \Gamma(z_1, z_2; F_\nu; \mathfrak{B}^4_{rr}) \tag{6.18}$$

are regular double harmonic functions in $\mathfrak{B}^4_{rr}$. Since

$$\Gamma(re^{i\varphi_1}, re^{i\varphi_2}; F_\nu; \mathfrak{B}^4_{rr}) = 0 \tag{6.19}$$

almost everywhere, we have by (2.12) that

$$\begin{aligned} &\log | F(0, 0) | + \sum_{\nu=1}^{N(r)} \Gamma(0, 0; F_\nu; \mathfrak{B}^4_{rr}) \\ &\qquad \leqq \left(\frac{1}{2\pi}\right)^2 \int_0^{2\pi} \int_0^{2\pi} \log^+ | F(re^{i\varphi_1}, re^{i\varphi_2}) | \, d\varphi_1 \, d\varphi_2 \leqq c_F. \end{aligned} \tag{6.20}$$

On the other hand, among the F_ν ($\nu = 1, \cdots, N(r)$) there occur the functions $F_1, \cdots, F_{p(r)}$ which vanish in $(\alpha^1), \cdots, (\alpha^{p(r)})$, i.e., in all the points $\{\alpha^\nu\}$ which lie in $\mathfrak{B}^4_{rr}$. Since $\Gamma(0, 0; F_\nu; \mathfrak{B}^4_{rr}) \geqq 0$ and

$$C(\alpha^\nu, r, r) \leqq \Gamma(0, 0; F_\nu, \mathfrak{B}^4_{rr}) \qquad (\nu = 1, \cdots, p), \tag{6.21}$$

we obtain

$$
\begin{aligned}
\sum_{\nu=1}^{p(r)} C(\alpha^{\nu}, r, r) &\leqq \sum_{\nu=1}^{p(r)} \Gamma(0, 0; F_{\nu}; \mathfrak{B}^{4}_{rr}) \\
&\leqq \sum_{\nu=1}^{N(r)} \Gamma(0, 0; F_{\nu}; \mathfrak{B}^{4}_{rr}) \leqq c_{F} - \log | F(0, 0) |.
\end{aligned} \tag{6.22}
$$

Since this inequality is true for every $r < 1$, it follows that

$$\sum_{\nu=1}^{\infty} C(\alpha^{\nu}) \leqq c_{F} - \log | F(0, 0) | < \infty. \tag{6.23}$$

This concludes the proof of Theorem 4.

Bibliography

S. Bergman.

1. *Über die Nullstellen einer Funktion von zwei komplexen Veränderlichen*, Koninklijke Akademie van Wetenschappen te Amsterdam, Proceedings of the Section of Sciences, vol. 35(1932), pp. 1188–1194.
2. *Sur les fonctions entières et méromorphes de deux variables complexes*, Compositio Mathematica, vol. 3(1936), pp. 136–173, and Recueil Mathématique, vol. 2(44) (1937), pp. 599–616.
3. *On a generalized Green's function and its application.* (To appear in the Bulletin of the American Mathematical Society.)

S. Bochner.

1. *Darstellung reell-variabler und analytischer Funktionen durch verallgemeinerte Fourier- und Laplace-Integrale*, Math. Annalen, vol. 97(1927), pp. 635–662.
2. *Bounded analytic functions in several variables and multiple Laplace integrals*, Amer. Jour. Math., vol. 59(1937), pp. 732–738.

C. H. Müntz.

1. *Über den Approximationssatz von Weierstrass*, Schwarz Festschrift, Berlin, 1914, pp. 303–312.

W. F. Osgood.

1. *Lehrbuch der Funktionentheorie*, vol. II, Berlin and Leipzig, 1929.

R. E. A. C. Paley and N. Wiener.

1. *Notes on the theory and application of Fourier transforms*, III, Trans. Amer. Math. Soc., vol. 35(1933), pp. 761–766.
2. *Fourier Transforms in the Complex Domain*, Amer. Math. Soc. Colloquium Publications, vol. 19, New York, 1934.

O. Szász.

1. *Über die Approximation stetiger Funktionen durch lineare Aggregate von Potenzen*, Mathematische Annalen, vol. 77(1916), pp. 482–496.

Massachusetts Institute of Technology.

GILLESPIE MEASURE

By Anthony P. Morse and John F. Randolph

1. **Remarks.** One aim of the theory of measure as introduced by Lebesgue was to furnish a tool for handling questions of length for curves and area for surfaces. While this theory was successful as far as curve length is concerned, the notion of area for surfaces was left in an unsatisfactory state. The fact that some of the properties of lengths for curves have not been successfully extended to areas for surfaces may suggest that our present notion of length for curves is perhaps of too special a nature. In fact even before Lebesgue had given his definition of the measure of a point set on a straight line, Minkowski [9][1] in 1909 had already considered the question of generalizing curve length by assigning (in the spirit of Peano-Jordan content) a linear measure to point sets lying in the plane.

This notion of assigning a linear measure to point sets not lying on a line has been considered in different ways since Minkowski with varying degrees of success. Among such definitions are those of Young 1905 [13], Janzen 1907 [7], Carathéodory 1914 [3], and Gross 1918 [5]. The Minkowski measure inherits all the faults of Peano-Jordan content. Young himself found inconsistencies for his own measure. As shown by Saks [11] Gross measure has the anomaly of assigning the measure zero to a particular set and a positive (even infinite) measure to the transform of this set by a bounded transformation of the plane. Janzen measure is not independent of the coördinate system, e.g., the set constructed by Gross ([5], p. 185) has Janzen measure unity for the axes used in the construction and measure zero when the axes are rotated 45°.

Until now Carathéodory measure has received practically no adverse criticism in the literature [8, 10] and seems to have been accepted as an adequate generalization of the notion of length. However, we give incidentally (§6) a set for which even this measure is quite inconsistent with our inherent concept of length.

2. **Introduction.** In this paper we propose still another definition of linear measure for point sets not necessarily lying on a line. For a point set A we shall represent this measure by $G^*(A)$ and call it Gillespie linear measure after the late Professor D. C. Gillespie who suggested to us individually definitions similar to the one we have adopted.

Gillespie outer linear measure is defined (§3) in such a way that Carathéodory's postulational theory of measure may be used.

Received November 13, 1939.

[1] Numbers in brackets refer to the bibliography at the end of the paper.

It will be obvious that Carathéodory linear measure is never greater than Gillespie linear measure:

$$L^*(A) \leqq G^*(A);$$

and it is easily shown that $G^*(A) \leqq \pi L^*(A)$. We prove, however, in §8 that

$$L^*(A) \leqq G^*(A) \leqq \frac{\pi}{2} L^*(A).$$

This relation is the best possible since there are sets for which each equality holds.

In §6 we establish a relation which seems indispensable to a generalization of Euclidean length; namely, if $| P_1 |$ and $| P_2 |$ are the outer Lebesgue measures of the projections of a plane set A on two perpendicular lines of the plane, then

$$G^*(A) \geqq (| P_1 |^2 + | P_2 |^2)^{\frac{1}{2}}.$$

The analogous relation for Carathéodory measure is not satisfied. In fact there is a set, see §5, with $L^*(A) = | P_1 | = | P_2 |$.

Also (§7) if a simple Jordan arc has length in the ordinary sense, then the set of all points of this arc has its Gillespie linear measure equal to the length of the arc.

Gillespie linear measure $G(A)$ is extended (§9) to Gillespie area measure $G^{(2)}(B)$ for sets B not necessarily lying in a plane. We prove that if A is a set in the (x, y)-plane and B is the set of all points (x, y, z) with (x, y) a point of A and $0 \leqq z \leqq h$, i.e.,

$$B = \underset{(x,y,z)}{\overset{*}{E}} [(x, y) \in A, 0 \leqq z \leqq h],$$

then $G^{(2)}(B) = hG(A)$. This relation between "area" and "length", fundamental and simple as it seems, has not been proved [10] for any of the other measures mentioned above.

3. **Definition of** $G^*(A)$. In this section A will designate a plane point set.

If U is a plane convex set with inner points, then by $c(U)$ we mean the length of the simple closed curve which is its boundary; if U is a segment, then $c(U)$ is twice the distance between its end points.

With ρ an arbitrary positive number let $U_1, U_2, \cdots$ be a sequence of convex sets lying in the plane each with diameter $< \rho$ whose union contains A, and consider the series of semi-circumferences

$$\sum \tfrac{1}{2}c(U_k).$$

Designate the greatest lower bound (which may be infinite) of all such numbers by $G_\rho(A)$. As ρ decreases, $G_\rho(A)$ does not decrease. Thus as $\rho \to 0$ the function $G_\rho(A)$ approaches a limit (finite or infinite) which is represented by

$$G^*(A)$$

and called the *Gillespie outer linear measure of* A.

If U is a convex set, its closure $\bar{U}$ is also convex and moreover $c(U) = c(\bar{U})$. Thus in the definition of $G^*(A)$ we could have restricted the sets $U_1, U_2, \cdots$ to be closed convex sets. Also, if only open convex sets $U_1, U_2, \cdots$ are used, the same number $G^*(A)$ is obtained. For with $\epsilon > 0$ arbitrary we may include U_k in an open convex set V_k with $c(U_k) \leqq c(V_k) \leqq c(U_k) + 2^{-k}\epsilon$.

One will see that Gillespie outer linear measure possesses the first four properties demanded of an outer measure function in Carathéodory's postulational development of the theory of measure [3, 4]. Thus if we accept Carathéodory's general definition of measurability,[2] it follows that the complements of measurable sets are measurable, that the intersection of a sequence of measurable sets is measurable, and that open sets are measurable. In particular, G_δ's, F_σ's, etc. are measurable. If a set A is measurable, its measure $G(A)$ is defined by the equation $G(A) = G^*(A)$.

Gillespie linear measure also satisfies Carathéodory's fifth axiom and, in addition, the following modification given by Hahn ([6], p. 444):

For each set A there is a set B which is a G_δ containing A such that $G(B) = G^*(A)$.

For, with $\rho_1, \rho_2, \cdots$ a decreasing sequence of numbers approaching zero, let $U_{k_1}, U_{k_2}, \cdots$ be open convex sets (just shown to exist) each with diameter $< \rho_k$ whose union contains A and such that $\sum_{n=1}^{\infty} \frac{1}{2}c(U_{k_n}) \leqq G^*(A) + \rho_k$. The set $B_k = U_{k_1} + U_{k_2} + \cdots$ is then open so $B = B_1 \cdot B_2 \cdots$ is a G_δ. Moreover, $B \supset A$ so $G(B) \geqq G^*(A)$. On the other hand, $B \subset U_{k_1} + U_{k_2} + \cdots$ and the diameter of U_{k_n} is $< \rho_k$, so $G_{\rho_k}(B) \leqq \sum_n \frac{1}{2}c(U_{k_n}) \leqq G^*(A) + \rho_k$ and thus $G(B) \leqq G^*(A)$.

The inner measure $G_*(A)$ of a set A is defined as the upper limit of the measures of all measurable subsets of A. From this definition and the above modification of the fifth axiom, Hahn ([6], p. 445) proves a theorem equivalent to the following statement:

If A has $G_*(A)$ finite, then this inner linear measure is the upper limit of the linear measures $G(K)$ of all closed subsets K of A.

Thus since a set with outer measure finite is measurable if and only if its outer and inner measures are the same ([4], p. 263), we have

THEOREM 1. *If $G^*(A) < \infty$, then a necessary and sufficient condition that A be Gillespie linearly measurable is that $\epsilon > 0$ imply the existence of a closed subset K of A with $G^*(A - K) < \epsilon$.*

4. **Relation of $L^*(A)$ and $G^*(A)$.** The definition in §3 of $G^*(A)$ is patterned closely after Carathéodory's definition [3] of the outer linear measure $L^*(A)$. In fact in Carathéodory's definition we have merely substituted the semicircumference $\frac{1}{2}c(U_k)$ where he has the diameter $d(U_k)$ of U_k.[3]

[2] I.e., a set A is *Gillespie linearly measurable* if for every set W with $G^*(W)$ finite the equality $G^*(W) = G^*(AW) + G^*(W - AW)$ holds.

[3] Carathéodory did not actually require the sets $U_1, U_2, \cdots$ to be convex, but showed that $L^*(A)$ is not altered if they are so restricted.

Thus, since $d(U_k) \leqq \frac{1}{2}c(U_k)$ and since $\frac{1}{2}c(U_k) \leqq \pi d(U_k)$ (because U_k may be included in a circle of radius $d(U_k)$), it follows that

$$(1) \qquad L^*(A) \leqq G^*(A) \leqq \pi L^*(A).$$

Consequently a set A has $L^*(A)$ and $G^*(A)$ both finite or both infinite and if either is zero the other is also.

Hahn's modification of Carathéodory's fifth axiom is also satisfied by $L^*(A)$ as one may see by following arguments similar to those used for $G^*(A)$. Thus the analogue of Theorem 1 holds for $L^*(A)$. Consequently we have

THEOREM 2. *If A has $G^*(A)$ (or $L^*(A)$) finite, then a necessary and sufficient condition for A to be Gillespie linearly measurable is that A be Carathéodory linearly measurable.*

For, from (1) a closed subset K of A exists with $G^*(A - K)$ arbitrarily small if and only if $L^*(A - K)$ is arbitrarily small.

5. A particular set. It might appear, from the similarity of the results in §4 for Gillespie and Carathéodory linear measures, that the difference between these two measures is so slight as to be of no significance. However, in this section we construct a set A whose projection on the x-axis is the closed segment $\underset{(x,y)}{E}\,[0 \leqq x \leqq 1,\ y = 0]$ and whose projection on the y-axis is $\underset{(x,y)}{E}\,[x = 0, 0 \leqq y \leqq 1]$, but nevertheless (instead of having $L^*(A)$ at least $2^{\frac{1}{2}}$) has $L^*(A) = 1$. In the next section we show that for $G^*(A)$ this irregularity cannot occur.

Toward constructing this set A we first define an operation of order n on a circle[4] C of radius r and with respect to a fundamental coördinate system.

Circumscribe a square about C with sides parallel to the axes. Draw a chord of C from the upper to the right-hand point of tangency and another chord from the left-hand to the lower point of tangency. Divide the square into $(2n)^2$ equal squares with sides parallel to the axes. Of the smaller squares thus formed consider *only* those which have a diagonal lying along one of the chords just drawn. From these squares select those which are subsets of C; those which are not entirely subsets of C redivide into n^2 equal squares. Of these still smaller squares which lie along either of the above chords, select those which are subsets of C and those which are not entirely subsets of C redivide into n^2 equal squares, etc. The selected squares are then infinite in number, no two overlap, each is a subset of C, and their union contains all except the end points of the two chords of C indicated above. In each of these squares inscribe a circle. The infinite set of circles thus obtained is to be considered the result of the operation of order n on C.

Notice in particular that for any n:

(1) The sum of the diameters of the circles obtained is $2r$.

(2) The union of the circles obtained may be included in two rectangles each of length $2^{\frac{1}{2}}r$ and width r/n.

[4] I.e., all points whose distance from a fixed point is $\leqq r$.

Start with a circle of diameter unity tangent to both axes in the first quadrant and perform the operation of order 3. Call the union of the circles thus obtained Γ_3. Then the closure $\bar{\Gamma}_3$ is Γ_3 together with the four points $(0, \frac{1}{2})$, $(\frac{1}{2}, 0)$, $(1, \frac{1}{2})$ and $(\frac{1}{2}, 1)$. On each circle of Γ_3 perform the operation of order 4 and call the union of all circles thus obtained Γ_4. Then $\bar{\Gamma}_4$ contains only a countable set of points more than Γ_4. In general, on each circle of Γ_{n-1} we perform the operation of order n and obtain a set $\Gamma_n \subset \Gamma_{n-1}$, and note that $\bar{\Gamma}_n - \Gamma_n$ is countable.

The set A we shall consider is the intersection

$$A = \bar{\Gamma}_3\bar{\Gamma}_4 \cdots \bar{\Gamma}_n \cdots = \lim_{n\to\infty} \bar{\Gamma}_n.$$

Thus A is a closed (and even perfect) set and is thus measurable.

For any non-negative constant $a \leqq 1$, the line $x = a$ (and also the line $y = a$) intersects $\bar{\Gamma}_n$ in a non-empty closed set so this line contains a point of A. Thus the projection of A on either axis is the unit interval $[0, 1]$.

Since the diameter of any set U is greater than or equal to the diameter of the projection of U on a line, it is seen that $L(A) \geqq 1$. On the other hand $\bar{\Gamma}_n$ is clearly covered by a countable number of convex sets each with diameter less than n^{-1}, such that the sum of the diameters is unity. Hence $L(A) \leqq 1$, and therefore $L(A) = 1$.

For later use in connection with the Gillespie linear measure of this particular set A we make a further observation. From property (2) above, the subset of A in any one circle of Γ_n may be included in two rectangles the sum of whose semi-circumferences is $(2^{\frac{1}{2}} + n^{-1})$ multiplied by the diameter (which is $\leqq (2n)^{-1}$) of that circle. Thus since $A - A\Gamma_n$ is countable, we may include all of A in a countable number of convex sets each with diameter less than n^{-1}, such that the sum of the semi-circumferences is not greater than $(2^{\frac{1}{2}} + n^{-1})$ multiplied by the sum (unity) of the diameters of the circles of Γ_n. Thus

$$G(A) \leqq 2^{\frac{1}{2}}.$$

6. **Projection properties.** Let P be a closed polygon with $c(P) = s_1 + \cdots + s_n$, where s_k is the distance between consecutive vertices of P, and let a_k be the distance between the projections of these two vertices on the x-axis and b_k on the y-axis. Let $| P^x |$ be the distance between the end points of the projection of P on the x-axis and $| P^y |$ for the y-axis. It follows that

$$c(P) = \sum_{k=1}^{n} (a_k^2 + b_k^2)^{\frac{1}{2}} \geqq [(\sum a_k)^2 + (\sum b_k)^2]^{\frac{1}{2}} \geqq [(2\,| P^x |)^2 + (2\,| P^y |)^2]^{\frac{1}{2}}.$$

Thus if U is any convex set,

$$c(U) \geqq 2(| U^x |^2 + | U^y |^2)^{\frac{1}{2}}. \tag{1}$$

We now prove the following theorem, the analogue of which is seen, by the example of §5, not to hold for Carathéodory measure.

THEOREM 3. *If A is any plane set, $m^*(A^x)$ and $m^*(A^y)$ the outer Lebesgue measures of the projections of A on the x- and y-axes respectively, then*[5]

$$G^*(A) \geqq \{[m^*(A^x)]^2 + [m^*(A^y)]^2\}^{\frac{1}{2}}.$$

For let $U_1, U_2, \cdots$ be a covering of A by convex sets such that $G^*(A) + \epsilon > \Sigma\frac{1}{2}c(U_k)$. From (1) this sum is $\geqq \Sigma(|U_k^x|^2 + |U_k^y|^2)^{\frac{1}{2}}$ which from Minkowski's inequality is $\geqq [(\Sigma|U_k^x|)^2 + (\Sigma|U_k^y|)^2]^{\frac{1}{2}}$. But ΣU_k^x is a covering of A^x so $\Sigma|U_k^x| \geqq m^*(A^x)$, and we see the desired result.

The set A constructed in §5 was seen to have

$$A^x = \underset{(x,y)}{E}[0 \leqq x \leqq 1, y = 0] \quad \text{and} \quad A^y = \underset{(x,y)}{E}[x = 0, 0 \leqq y \leqq 1]$$

and thus from Theorem 3 to have $G(A) \geqq 2^{\frac{1}{2}}$. But at the end of §5 we saw that $G(A) \leqq 2^{\frac{1}{2}}$, so for this particular set we have the exact relations

$$m(A^x) = m(A^y) = L(A) = 1, \qquad G(A) = 2^{\frac{1}{2}}.$$

7. **Length of a curve.** Let γ be a curve without double points, i.e., a unique image of the closed interval $0 \leqq t \leqq 1$ by continuous functions φ and ψ such that $\varphi(t) = \varphi(t')$ and $\psi(t) = \psi(t')$ implies $t = t'$. We shall let (γ) represent the set of points on the curve γ. The set (γ) is then closed and thus Gillespie linearly measurable.

In this section we shall prove

THEOREM 4. *The Gillespie linear measure $G(\gamma)$ of the set (γ) is the upper limit λ of numbers of the form*

$$\sum_{i=1}^{n} \{[\varphi(t_i) - \varphi(t_{i-1})]^2 + [\psi(t_i) - \psi(t_{i-1})]^2\}^{\frac{1}{2}}, \tag{1}$$

where $0 = t_0 < t_1 < \cdots < t_n = 1$, i.e., $G(\gamma)$ is the length of the curve γ.

If $0 = t_0 < \cdots < t_n = 1$ and γ_i is the part of γ corresponding to the open interval (t_{i-1}, t_i), then by Theorem 3

$$G(\gamma_i) \geqq \{[\varphi(t_{i-1}) - \varphi(t_i)]^2 + [\psi(t_{i-1}) - \psi(t_i)]^2\}^{\frac{1}{2}} \qquad (i = 1, 2, \cdots, n);$$

hence

$$G(\gamma) \geqq \sum_{i=1}^{n} G(\gamma_i) \geqq \sum_{i=1}^{n} \{[\varphi(t_{i-1}) - \varphi(t_i)]^2 + [\psi(t_{i-1}) - \psi(t_i)]^2\}^{\frac{1}{2}}$$

[5] For Janzen measure $J^*(A)$ one will sometimes see the relation

$$J^*(A) \geqq \{[m^*(A^x)]^2 + [m^*(A^y)]^2\}^{\frac{1}{2}}$$

(e.g., [12], p. 5). However, Janzen measure is defined with respect to and may not be independent of the coördinate system, so in this relation it is mandatory that the x and y refer to the coördinate system with respect to which $J^*(A)$ is defined. On the other hand, $G^*(A)$ is independent of the coördinate system.

and

$$G(\gamma) \geqq \lambda. \tag{2}$$

We now prove the reverse inequality.

Since λ is the upper limit of numbers of the form (1), we have, for $0 \leqq t_0 \leqq 1$,

$$\{[\varphi(0) - \varphi(t_0)]^2 + [\psi(0) - \psi(t_0)]^2\}^{\frac{1}{2}} + \{[\varphi(t_0) - \varphi(1)]^2 + [\psi(t_0) - \psi(1)]^2\}^{\frac{1}{2}} \leqq \lambda,$$

i.e., the set (γ) is included in an ellipse with one axis of length λ and (d denoting the distance between the end points of γ) the other axis of length $(\lambda^2 - d^2)^{\frac{1}{2}}$. We may thus include the set (γ) in a rectangle with dimensions λ and $(\lambda^2 - d^2)^{\frac{1}{2}}$.

Let $\rho > 0$ be an arbitrary number. Choose $0 = t_0 < t_1 < \cdots < t_n = 1$ such that if d_i is the distance between the points $(\varphi(t_{i-1}), \psi(t_{i-1}))$ and $(\varphi(t_i), \psi(t_i))$ and λ_i is the length of the arc γ_i of γ joining these two points, then

$$\lambda_i < \tfrac{1}{2}\rho \quad \text{and} \quad [\lambda^2 - (\sum_{i=1}^{n} d_i)^2]^{\frac{1}{2}} < \rho.$$

Now enclose (γ_i) in a rectangle U_i of dimensions λ_i and $(\lambda_i^2 - d_i^2)^{\frac{1}{2}}$ as we have just shown is possible. The rectangle U_i has diagonal $< \rho$ and is a convex set, so we have[6]

$$G_\rho(\gamma) \leqq \sum_{i=1}^{n} \tfrac{1}{2}c(U_i) = \sum_{i=1}^{n} \lambda_i + \sum_{i=1}^{n} (\lambda_i^2 - d_i^2)^{\frac{1}{2}} \leqq \lambda + [\lambda^2 - (\sum_{i=1}^{n} d_i)^2]^{\frac{1}{2}} \leqq \lambda + \rho.$$

Since ρ is arbitrary, we thus have

$$G(\gamma) \leqq \lambda,$$

the inequality reverse from (2).

8. $G^*(A) \leqq \frac{1}{2}\pi L^*(A)$. From the fact that a point set of diameter d may be included in a circle of diameter $2d$, i.e., semi-circumference πd, we have already seen that $G^*(A) \leqq \pi L^*(A)$. A theorem by Young ([1], p. 463) states that a set of diameter d may be included in a circle of diameter $2 \cdot 3^{-\frac{1}{2}}d$ and thus the sharper inequality $G^*(A) \leqq 3^{-\frac{1}{2}}\pi L^*(A)$ may be obtained. In this section we prove still more; we prove

THEOREM 5. *If A is an arbitrary plane point set, then*

$$G^*(A) \leqq \tfrac{1}{2}\pi L^*(A).$$

With $\rho > 0$, let $U_1, U_2, \cdots$ be a covering of A by convex sets such that the diameter $d(U_i) < \rho$ and $\Sigma d(U_i) \leqq L^*(A) + \rho$. Then $G_\rho(A) \leqq \Sigma\frac{1}{2}c(U_i)$.

[6] This inequality follows since $0 \leqq a_i \leqq b_i$ implies

$$\sum (b_i^2 - a_i^2)^{\frac{1}{2}} \leqq [(\sum b_i)^2 - (\sum a_i)^2]^{\frac{1}{2}}$$

and because $\sum_{i=1}^{n} \lambda_i \leqq \lambda$.

But (see [2], p. 65) a plane convex set U of diameter d has $c(U) \leqq \pi d$. Thus $G_\rho(A) < \frac{1}{2}\pi[L^*(A) + \rho]$ and since ρ is arbitrary, we have the desired result.

Furthermore, this result is the best possible, for there are sets for which the equality holds. A set constructed for a different purpose by Besicovitch ([1], p. 431) is one such set, but the proof will not be given.

9. **Gillespie area measure.** Let A be a point set in 3-dimensional space.

The surface area $s(U)$ of a convex set U in 3-dimensional space is defined as inf H, where H is the set of real numbers determined by: $p \in H$ if there exist an open convex set P and closed triangles $\Delta_1, \Delta_2, \cdots, \Delta_n$ for which

$$U \subset P, \qquad \text{boundary of } P \subset \sum_{j=1}^{n} \Delta_j, \qquad p = \sum_{j=1}^{n} (\text{area of } \Delta_j).$$

The definition of Gillespie outer area measure $G^{*2}(A)$ is obtained from the definition of Gillespie outer linear measure (§3) by replacing the words "the plane" by "3-dimensional space" and $c(U_k)$ by $s(U_k)$.

Furthermore the statements made in §3 about Gillespie linear measure all have analogues for Gillespie area measure.

10. **Relation between Gillespie linear and area measures.** In the Lebesgue theory of integration if M is a bounded Lebesgue measurable set on the x-axis and f is a non-negative bounded Lebesgue measurable function on M, then the plane set $N_f = \underset{(x,y)}{E}\, [x \in M,\, 0 \leqq y \leqq f(x)]$ is Lebesgue plane measurable and $m^{(2)}(N_f) = \int_M f(x)\, dx$. This relation is equivalent to the fact that if f is a constant k, then the set N_f is Lebesgue plane measurable with

$$(1) \qquad m^{(2)}(N_f) = km(M).$$

To each of the linear measures mentioned in the introduction corresponds an analogous area measure for sets not necessarily lying in the plane. It has not been shown that any of these measures preserve, as does Lebesgue measure, the Euclidean relation that area is the product of length by length.

In this section we show, however, that if A is a set in the (x, y)-plane and B is the cylindrical set

$$B = \underset{(x,y,z)}{E}\, [(x, y) \in A,\, 0 \leqq z \leqq h],$$

then the Gillespie area of B is the Gillespie length of A multiplied by h.

Throughout this section A and B will be used to designate the sets given here.

We first prove, using $| U |$ for the Lebesgue 2-dimensional measure of the plane convex set U,

Lemma 1. *If $G^*(A)$ is finite, $\rho_1 > \rho_2 > \cdots, \rho_n \to 0$, and $U_{n1}, U_{n2}, \cdots$ a sequence of coverings of A by plane convex sets with $d(U_{nk}) < \rho_n$ such that $\lim_{n\to\infty} \sum_k \frac{1}{2}c(U_{nk}) = G^*(A)$, then $\lim_{n\to\infty} \sum_k | U_{nk} | = 0$.*

For U_{nk} may be included in a circle of radius $d(U_{nk})$ and thus

$$\sum_k | U_{nk} | \leqq \sum_k \pi\, d(U_{nk})^2 < \pi\rho_n \sum d(U_{nk}) \leqq \pi\rho_n \sum_k \tfrac{1}{2}c(U_{nk})$$

which approaches zero with n^{-1}.

We now obtain an inequality for outer measures which needs proof only if $G^*(A) < \infty$.

THEOREM 6. $G^{*(2)}(B) \leqq hG^*(A)$.

Given $\rho > 0$, take N an integer so large that $h/N < 2^{-\frac{1}{2}}\rho$. Then from the lemma we may cover A by convex sets $U_1, U_2, \cdots$ with $d(U_k) < 2^{-\frac{1}{2}}\rho$, $\Sigma\frac{1}{2}c(U_k) < G^*(A) + \rho/h$ and $\Sigma \mid U_k \mid < \rho/N$. Now let

$$V_{kn} = \underset{(x,y,z)}{E} [(x, y) \in U_k, nh/N \leqq z \leqq (n+1)h/N].$$

Then $d(V_{kn}) < \rho$ and $\sum_k \sum_{n=0}^{N} V_{kn} \supset B$. Thus

$$G_\rho^{(2)}(B) \leqq \sum_k \sum_{n=0}^{N} \frac{1}{2} s(V_{kn}) \leqq \sum_k \frac{2N \mid U_k \mid}{2} + N \cdot \frac{h}{N} \cdot \frac{1}{2} c(U_k)$$

$$\leqq N(\rho/N) + h[G^*(A) + \rho/h] = hG^*(A) + 2\rho.$$

Therefore since ρ is arbitrary, the theorem holds.

In particular, we have the

COROLLARY. *If A is Gillespie linearly measurable with $G(A) < \infty$, then B is Gillespie area measurable and $G^{(2)}(B) \leqq hG(A)$.*

For, from Theorem 1, $\epsilon > 0$ implies that a closed subset K of A exists with $G(A - K) < \epsilon$. Thus the set

$$H = \underset{(x,y,z)}{E} [(x, y) \in K, 0 \leqq z \leqq h]$$

is a closed subset of B such that, from (2), $G^{*(2)}(B - H) \leqq h\epsilon$. Thus B is Gillespie area measurable since, as one will see, the area analogue of Theorem 1 holds.

If A is a set in 3-dimensional space, we use the notation

$$A^z = \underset{(x,y)}{E} [(x, y, z) \in A].$$

LEMMA 2. *If Δ is a closed triangle in 3-dimensional space, then*

$$\int_{-\infty}^{\infty} G(\Delta^z)\, dz \leqq \text{area of } \Delta.$$

LEMMA 3. *If W is a bounded convex set in 3-dimensional space, then*

$$\int_{-\infty}^{\infty} c(W^z)\, dz \leqq s(W).$$

Let $\epsilon > 0$ be arbitrary, and let P be an open convex set with boundary P^* such that

(i) P is contained in W;

(ii) there exist closed triangles $\Delta_1, \Delta_2, \cdots, \Delta_n$ for which

$$P^* \subset \sum_{j=1}^{n} \Delta_j, \qquad \sum_{j=1}^{n} (\text{area of } \Delta_j) \leqq s(W) + \epsilon.$$

Now W^z and P^z are convex, $W^z \subset P^z$ and the boundary of P^z is contained in $\sum_{j=1}^{n} \Delta_j^z$; so we have[7]

$$c(W^z) \leqq c(P^z) \leqq \sum_{j=1}^{n} G(\Delta_j^z).$$

Hence, using the previous lemma, we obtain

$$\int_{-\infty}^{\infty} c(W^z)\,dz \leqq \sum_{j=1}^{n} \int_{-\infty}^{\infty} G(\Delta_j^z)\,dz \leqq \sum_{j=1}^{n} (\text{area of } \Delta_j) \leqq s(W) + \epsilon.$$

The proof is complete.

THEOREM 7. *If A has $G_*(A) < \infty$, then $G_*^{(2)}(B) \geqq hG_*(A)$.*

First let K be a closed subset of A with $G(K) < \infty$. Then the set $H = \underset{(x,y,z)}{E} [(x, y) \epsilon K, 0 \leqq z \leqq h]$ is a closed subset of B with, from Theorem 6, $G^{(2)}(H) < \infty$.

From the definition of Gillespie outer linear measure, corresponding to an $\epsilon > 0$ there exists a $\rho > 0$ such that $G(K) - \epsilon < \Sigma\frac{1}{2}c(U_k)$ if $U_1, U_2, \cdots$ is a covering of K by plane convex sets with $d(U_k) < \rho$. We then choose a covering $W_1, W_2, \cdots, W_n$ of H by three-dimensional open convex sets with $d(W_k) < \rho$ such that

$$G^{(2)}(H) + \epsilon > \sum_{k=1}^{n} \frac{1}{2} s(W_k);$$

the numbers of the covering reduced to being finite in number since H is closed. Then W_k^z is convex with diameter $< \rho$ and, for $0 \leqq z \leqq h$, the union $\sum_{k=1}^{n} W_k^z$ contains the plane set H^z which is the set K. Thus

$$\sum_{k=1}^{n} \frac{1}{2} c(W_k^z) > G(K) - \epsilon, \qquad 0 \leqq z \leqq h.$$

But, from Lemma 3,

$$s(W_k) \geqq \int_{-\infty}^{\infty} c(W_k^z)\,dz \geqq \int_0^h c(W_k^z)\,dz.$$

[7] [2], p. 47.

Also since we have a finite number of terms,

$$\sum_{k=1}^{n} \int^{h} \frac{1}{2} c(W_k^z)\, dz = \int_0^h \left[\sum_{k=1}^{n} \frac{1}{2} c(W_k^z)\right] dz.$$

Thus $G^{(2)}(H) + \epsilon > \int_0^h [G(K) - \epsilon]\, dz$, and since ϵ is arbitrary,

$$G^{(2)}(H) \geqq hG(K). \tag{2}$$

Now since $B \supset H$, we have $G_*^{(2)}(B) \geqq G^{(2)}(H)$. Since $G_*(A)$ is finite, $G(K)$ may be taken arbitrarily close to it and the desired result $G_*^{(2)}(B) \geqq hG_*(A)$ holds.

From the corollary to Theorem 6 and Theorem 7 we have

COROLLARY 1. *If A is Gillespie linearly measurable with $G(A)$ finite, then B is Gillespie arealy measurable and $G^{(2)}(B) = hG(A)$.*

We may now prove

THEOREM 8. *If A is an arbitrary plane set with $G^*(A) < \infty$, then*

$$G_*^{(2)}(B) = hG_*(A) \quad \textit{and} \quad G^{*(2)}(B) = hG^*(A).$$

With $G_*(A) < \infty$, let H' be a bounded closed subset of B with $G_*^{(2)}(B) - \epsilon < G^{(2)}(H')$, such a set existing from the definition of inner measure and the analogue for area measure of Theorem 2. Being bounded and closed, the set $K = \underset{(x,y)}{E} [(x, y, z) \epsilon H']$ is also closed. Thus the set

$$H = \underset{(x,y,z)}{E} [(x, y) \epsilon K, 0 \leqq z \leqq h]$$

contains H, so that $G^{(2)}(H') \leqq G^{(2)}(H)$; and since H is measurable (closed), $G^{(2)}(H) = hG(K)$ from Corollary 1. But $H' \subset B$ so $K \subset A$ and $G(K) \leqq G_*(A)$. Since ϵ is arbitrary,

$$G_*^{(2)}(B) \leqq hG_*(A),$$

and this with Theorem 7 gives the desired equality for inner measure.

For outer measures, let $\tilde{A}$ be a measurable set containing A with $G(\tilde{A}) < \infty$. Then (see [4], p. 262, Theorem 4)

$$G(\tilde{A}) = G^*(A) + G_*(\tilde{A} - A).$$

Then multiplying by h and using Corollary 1 and the first equality of the corollary under proof, we have

$$G^{(2)}(\tilde{B}) = hG^*(A) + G_*^{(2)}(\tilde{B} - B),$$

where $\tilde{B} = \underset{(x,y,z)}{E} [(x, y) \epsilon \tilde{A}, 0 \leqq z \leqq h]$. But $G^{(2)}(\tilde{B}) < \infty$ and

$$G^{(2)}(\tilde{B}) = G^{*(2)}(B) + G_*^{(2)}(\tilde{B} - B),$$

so we have the desired equality

$$G^{*(2)}(B) = hG^*(A).$$

In 3-dimensional space a set which is arealy measurable need not project orthogonally on a plane into a linearly measurable set. However, with A and B the related sets of this section, the converse of Corollary 1 holds, i.e., we have

COROLLARY 3. *If B is a Gillespie arealy measurable set with $G^{(2)}(B)$ finite, then A is Gillespie linearly measurable with*

$$G^{(2)}(B) = hG(A).$$

For $G^{*(2)}(B) = G_*^{(2)}(B)$. From Theorem 8, $G^*(A) = G_*(A)$ and the desired equality holds.

BIBLIOGRAPHY

1. A. S. BESICOVITCH, *On the fundamental geometrical properties of linearly measurable sets of points*, Mathematische Annalen, vol. 98(1928), pp. 422–464.
2. T. BONNESEN AND W. FENCHEL, *Theorie der konvexen Körper*, Ergebnisse der Mathematik, vol. 3(1935), pp. 1–164.
3. C. CARATHÉODORY, *Über das lineare Mass von Punktmengen eine Verallgemeinerung des Längenbegriffs*, Göttinger Nachrichten, (1914), pp. 404–426.
4. C. CARATHÉODORY, *Vorlesungen über Reelle Funktionen*, 1927.
5. W. GROSS, *Über Flächenmass von Punktmengen*, Monatshefte für Mathematik und Physik, vol. 29(1918), pp. 145–176; and *Über das lineare Mass von Punktmengen*, same volume, pp. 177–193.
6. H. HAHN, *Theorie der reellen Funktionen*, vol. 1, 1921.
7. O. JANZEN, *Über einige stetige Kurven, über Bogenlänge, linearen Inhalt und Flächeninhalt*, Inaugural-Dissertation, Königsberg, 1907.
8. R. L. JEFFERY, *Sets of k-extent in n-dimensional space*, Transactions of the American Mathematical Society, vol. 35(1933), pp. 629–647.
9. H. MINKOWSKI, *Über die Begriffe Länge, Oberfläche und Volumen*, Jahresbericht der Deutschen Mathematiker-Vereinigung, vol. 9(1900), pp. 115–121.
10. J. F. RANDOLPH, *On generalizations of length and area*, Bulletin of the American Mathematical Society, vol. 42(1936), pp. 268–274.
11. S. SAKS, *Remarque sur la mesure linéaire des ensembles plans*, Fundamenta Mathematicae, vol. 9(1926), pp. 16–24.
12. J. P. SCHAUDER, *The theory of surface measure*, Fundamenta Mathematicae, vol. 8(1926), pp. 1–48.
13. W. H. YOUNG AND GRACE CHISHOLM YOUNG, *The Theory of Sets of Points*, 1906.

INSTITUTE FOR ADVANCED STUDY.

NOTE ON THE INVERSION OF THE LAPLACE INTEGRAL

By Harry Pollard

W. Feller and M. J. Dubourdieu have recently obtained the following simple inversion of a Laplace integral.[1]

If

$$f(x) = \int_0^\infty e^{-xt}\, d\alpha(t) \tag{1}$$

converges for $x > c$ with $\alpha(t)$ normalized[2] and non-decreasing in every finite interval, then

$$\alpha(t) = \lim_{x\to\infty} \sum_{n=0}^{[xt]} \frac{(-x)^n}{n!} f^{(n)}(x) \qquad (t > 0). \tag{2}$$

By an obvious change of variable we can write this in the form

$$\alpha(t) = \lim_{k\to\infty} \sum_{n=0}^{k} \frac{(-1)^n}{n!} \left(\frac{k+\theta_k}{t}\right)^n f^{(n)}\left(\frac{k+\theta_k}{t}\right) \qquad (t > 0), \tag{2'}$$

where $\{\theta_k\}$ is a sequence[3] satisfying $0 \leqq \theta_k < 1$.

Earlier, however, Widder had obtained an inversion of (1) on the weaker hypothesis that $\alpha(t)$ is normalized and of bounded variation on every finite interval. His conclusion was[4]

$$\alpha(t) = \lim_{k\to\infty} \sum_{n=0}^{k} \frac{(-1)^n}{n!} \left(\frac{k}{t}\right)^n f^{(n)}\left(\frac{k}{t}\right) \qquad (t > 0). \tag{2''}$$

By a comparison of (2′) and (2″) we are led to conjecture that the conclusion of Feller and Dubourdieu is not the best possible and that actually their operator (2) has the same degree of generality as the other. By methods closely related to Widder's we are able to establish this and even more, namely, that with the

Received November 23, 1939.

[1] W. Feller, *Completely monotone functions and sequences*, this Journal, vol. 5(1939), pp. 661–674; pp. 662–663.

M. J. Dubourdieu, *Sur un théorème de M. S. Bernstein relatif à la transformation de Laplace-Stieltjes*, Compositio Mathematica, vol. 7(1939), pp. 96–111.

These two authors obtained the result independently, Feller stating the equation in (2) only for points of continuity of $\alpha(t)$.

[2] I.e., $\alpha(0) = 0$, $\alpha(t) = \frac{1}{2}[\alpha(t+) + \alpha(t-)]$ $(t > 0)$.

[3] The sequence depends, of course, upon the particular value of t under consideration. In fact $\theta_k = xt - k$.

[4] D. V. Widder, *The inversion of the Laplace integral and the related moment problem*, Transactions of the American Mathematical Society, vol. 36(1934), pp. 107–200. We shall refer to this paper as W. The particular result to which we make reference here is Theorem 2, p. 116. The form of the conclusion found there is only apparently different from (2″), as an examination of the proof will disclose.

more general hypothesis on $\alpha(t)$ equation (2′) holds with $\{\theta_k\}$ now an *arbitrary* sequence satisfying $0 \leqq \theta_k \leqq 1$. The precise form of the extension is given by our Theorem 3.

In establishing this result we at the same time generalize some inversion theorems of Widder and Post for integrals of the form[5]

$$f(x) = \int_0^\infty e^{-xt}\phi(t)\, dt. \tag{3}$$

These generalizations constitute our first two theorems.

Use will be made of the following operator:[6] if $f(x)$ belongs to C^k, then

$$L_{k,t}[f(x)] = \frac{(-1)^k}{k!} x^{k+1} f^{(k)}(x) \,|_{x=(k+\theta_k)/t}, \tag{4}$$

where $\{\theta_k\}$ is a real sequence such that $0 \leqq \theta_k \leqq 1$ for all k.

By a change of variable of integration it is easily seen that in the special case that $f(x)$ is a function of the form (3) the operator becomes

$$L_{k,t}[f(x)] = \left(1 + \frac{\theta_k}{k}\right)^{k+1} \frac{k^{k+1}}{k!} \int_0^\infty e^{-(k+\theta_k)u} u^k \phi(tu)\, du. \tag{5}$$

We proceed to a group of lemmas which culminate in the results mentioned above. In all cases $\{\theta_k\}$ is the sequence defined in (4).

Lemma 1. *We have*

$$\left(1 + \frac{\theta_k}{k}\right)^k \sim e^{\theta_k} \qquad (k \to \infty). \tag{6}$$

The proof is easily supplied.

Lemma 2. *Let $f(x)$ have the representation* (1) *for $x > c$, where $\alpha(t)$ is of bounded variation on every finite interval. Then*

$$|\alpha(t)| < Me^{gt} \qquad (t \geqq 0),$$

where M and g are positive constants.[7]

Lemma 3. *Let $f(x)$ have the representation* (3) *for $x > c$, where $\phi(t)$ is integrable on every finite interval. Then*

$$\lim_{k\to\infty} \frac{k^{k+1}}{k!} \int_0^\infty e^{-ku} u^k \phi(tu)\, du = \phi(t) \tag{7}$$

for almost all $t > 0$. In particular (7) *holds at all points of continuity of $\phi(t)$.*[8]

[5] Real inversion methods, on the hypothesis that $\phi(t)$ is continuous over its entire domain, were first established by E. L. Post, *Generalized differentiation*, Transactions of the American Mathematical Society, vol. 32(1930), pp. 723–781; p. 772.

[6] Cf. W, p. 117. Observe that the operator given there reduces to ours in the case that all the θ_k are zero.

[7] D. V. Widder, *A generalization of Dirichlet's series and of Laplace's integrals by means of a Stieltjes integral*, Transactions of the American Mathematical Society, vol. 31(1929), pp. 694–743; Lemma 2, p. 703.

[8] W, Theorem 4, and Corollary, pp. 122–125.

If $\phi(t)$ is normalized and of bounded variation on every finite interval, then (7) *holds for all positive t.*[9]

This lemma summarizes those results of Widder which are generalized in Theorems 1 and 2.

THEOREM 1. *Let $f(x)$ have the representation* (3) *for $x > c$, where $\phi(t)$ is integrable on every finite interval. Then*

$$\lim_{k\to\infty} L_{k,t}[f(x)] = \phi(t) \tag{8}$$

for almost all $t > 0$. In particular, (8) *holds at all points of continuity of $\phi(t)$.*

We observe that this reduces to Lemma 3 in case all the θ_k of (4) are zero. Then it is enough to show that the difference between the operators (4) in the two cases approaches zero with $1/k$. This, by (5) and (6), amounts to showing that

$$I(k) = \frac{k^{k+1}}{k!}\int_0^\infty e^{-ku}u^k\{1 - e^{(1-u)\theta_k}\}\,\phi(tu)\,du = o(1) \qquad (k\to\infty) \tag{9}$$

for almost all $t > 0$, and in particular the points of continuity of $\phi(t)$.

Let $\epsilon > 0$ be arbitrary. Since $1 - e^{(1-u)\theta_k}$ approaches zero uniformly in k as $u \to 1$, there exists a δ $(0 < \delta < 1)$ such that

$$|\,1 - e^{(1-u)\theta_k}\,| < \epsilon \qquad (|\,1-u\,| < \delta;\ k = 0, 1, 2, \cdots). \tag{10}$$

Divide the interval of integration in (9) into the three parts $(0, 1-\delta)$, $(1-\delta, 1+\delta)$, $(1+\delta, \infty)$, and denote the corresponding integrals by I_1, I_2, I_3.

Since $e^{-u}u$ is increasing in $(0, 1)$, we have

$$|\,I_1\,| \leqq \frac{k^{k+1}}{k!}(1+e)e^{-k(1-\delta)}(1-\delta)^k\int_0^{1-\delta}|\,\phi(tu)\,|\,du,$$

and so

$$I_1 = o(1) \qquad (k\to\infty;\ t > 0). \tag{11}$$

Now let $\alpha_t(u) = \int_0^u \phi(tv)\,dv$. Then by Lemma 2, N and c_1, both numbers depending on t, exist such that

$$|\,\alpha_t(u)\,| < Ne^{c_1u} \qquad (u \geqq 0). \tag{12}$$

I_3 now becomes

$$I_3 = \frac{k^{k+1}}{k!}\int_{1+\delta}^\infty e^{-ku}u^k\{1 - e^{(1-u)\theta_k}\}\,d\alpha_t(u).$$

[9] W, Theorem 5, p. 126.

Taking $k > c_1$, and integrating by parts, we have

$$I_3 = -\frac{k^{k+1}}{k!} e^{-k(1+\delta)}(1+\delta)^k(1 - e^{-\delta\theta_k})\alpha_t(1+\delta)$$

$$- \frac{k^{k+1}}{(k-1)!} \int_{1+\delta}^{\infty} \alpha_t(u) e^{-ku} u^{k-1}\{(1 - e^{(1-u)\theta_k})(1-u) + \frac{\theta_k}{k} e^{(1-u)\theta_k} u\}\, du$$

$$= J_1 + J_2, \text{ say.}$$

Then J_1 approaches zero with $1/k$. Also by (12)

$$| J_2 | \leqq \frac{k^{k+1}}{(k-1)!} N \int_{1+\delta}^{\infty} e^{c_1 u} u^{k-1} e^{-ku}\{c_2 u\}\, du \qquad (k > c_1),$$

where c_2 is a positive constant. Then taking $\lambda > c_1$, we get

$$| J_2 | \leqq \frac{k^{k+1}}{(k-1)!} N c_2 (1+\delta)^{k-\lambda} e^{-(k-\lambda)(1+\delta)} \int_0^{\infty} e^{-(\lambda - c_1)u} u^{\lambda}\, du,$$

since $e^{-u}u$ is decreasing for $u > 1$. Hence J_2 also approaches zero. Then

$$I_3 = o(1) \qquad (k \to \infty;\ t > 0). \tag{13}$$

It remains to consider I_2. From (10) it follows that

$$| I_2 | < \frac{k^{k+1}}{k!} \epsilon \int_{1-\delta}^{1+\delta} e^{-ku} u^k \,| \phi(tu) |\, du = \epsilon \frac{k^{k+1}}{k!} \int_0^{\infty} e^{-ku} u^k \psi(tu)\, du \qquad (t > 0),$$

where $\psi(u) = | \phi(u) |$ in $(t - \delta t, t + \delta t)$ and zero otherwise. Let $k \to \infty$. Applying Lemma 3 to $\psi(u)$, we deduce that

$$\overline{\lim_{k \to \infty}}\, | I_2 | \leqq \epsilon \psi(t) = \epsilon \,| \phi(t) | \tag{14}$$

for almost all $t > 0$, and in particular the points of continuity of $\phi(t)$. From (11), (13), (14) it follows that

$$\overline{\lim_{k \to \infty}}\, | I(k) | \leqq \epsilon \,| \phi(t) | \tag{15}$$

at least on this set of t, and since ϵ is arbitrary, (9) follows.

Suppose now that in the preceding work $\phi(t)$ is normalized and of bounded variation in every finite interval. Then by the second part of Lemma 3, (14), and hence (15), holds for all $t > 0$. By the same argument as in the preceding theorem we thus establish

THEOREM 2. *Let $f(x)$ have the representation* (3) *for $x > c$, where $\phi(t)$ is normalized and of bounded variation on every finite interval. Then* (8) *holds for all $t > 0$.*

We are now in a position to establish our extension of the result of Feller and Dubourdieu.

Theorem 3. *Let (1) converge for $x > c$, where $\alpha(t)$ is normalized and of bounded variation on every finite interval, and let $\{\theta_k\}$ be any sequence such that $0 \leqq \theta_k \leqq 1$. Then (2′), and hence (2), is valid.*

Integrating (1) by parts and applying Lemma 2, we have $f(x) = xF(x)$, where

$$F(x) = \int_0^\infty e^{-xt}\alpha(t)\,dt \qquad (x > g). \tag{16}$$

Then

$$\begin{aligned}\sum_{n=0}^{k}\frac{(-1)^n}{n!}\left(\frac{k+\theta_k}{t}\right)^n f^{(n)}\left(\frac{k+\theta_k}{t}\right) &= -\sum_{n=0}^{k}\frac{(-1)^{n+1}}{n!}\left(\frac{k+\theta_k}{t}\right)^{n+1}F^{(n)}\left(\frac{k+\theta_k}{t}\right)\\ &\quad + \sum_{n=1}^{k}\frac{(-1)^n}{(n-1)!}\left(\frac{k+\theta_k}{t}\right)^n F^{(n-1)}\left(\frac{k+\theta_k}{t}\right)\\ &= \frac{(-1)^k}{k!}\left(\frac{k+\theta_k}{t}\right)^{k+1}F^{(k)}\left(\frac{k+\theta_k}{t}\right) = L_{k,t}[F(x)].\end{aligned}$$

Now let $k \to \infty$. By an application of Theorem 2 to $F(x)$ as given by (16), (2′) follows.

Harvard University.

THE JACOBI CONDITION FOR EXTREMALOIDS

By M. F. Smiley

In a recent paper McShane discussed the Jacobi condition for extremals in a manner which permitted the simultaneous consideration of the parametric and non-parametric problems.[1] It is the purpose of this note to formulate the Jacobi condition for extremaloids in a similar fashion. In order to obtain sufficient flexibility at corners we find it desirable to alter McShane's (Bliss') definition of normality of solutions of the Jacobi equations.[2] No attempt is made to discuss questions of tensor invariance.[3]

A knowledge of the details of McShane's paper, particularly §§2, 4, 5, and 11, is presupposed.

We begin with the consideration of a non-singular extremaloid g: $\gamma^i(t)$ $(i = 1, \cdots, n; t_1 \leqq t \leqq t_2)$ with corners at $t = x_\theta$ $(\theta = 1, \cdots, r)$ at which $\Omega_0 \neq 0$, where[4]

$$\Omega_0(t) = f_{y^i}(t^+)\dot{\gamma}^i(t^-) - f_{y^i}(t^-)\dot{\gamma}^i(t^+).$$

In our formulation of the Jacobi condition we shall use *accessory pseudo-extremaloids* which consist of functions $u^i(t)$, $\tau(t)$ $(t_1 \leqq t \leqq t_2)$ with the following properties:

(1) The functions $u^i(t)$, $\tau(t)$ are of class C^2 between corners[5] of g.

(2) The functions $u^i(t)$ satisfy the Jacobi equations between corners of g.

(3) The functions $u^i + \tau\dot{\gamma}^i$, $\Omega_{\rho^i}(u, \dot{u}) + \tau f_{y^i}$, τ are continuous on the whole interval t_1t_2.

We shall see that the values of the function $\tau(t)$ between corners of g are unimportant.

We shall suppose that there is a vector $p(t)$ with the properties:

(1) Each $p_i(t)$ is of class C^2 between corners of g.

(2) $p_i(t)\dot{\gamma}^i(t) \neq 0$ $(t_1 \leqq t \leqq t_2)$.

Received December 2, 1939.

[1] E. J. McShane, *The Jacobi condition and the index theorem in the calculus of variations*, this Journal, vol. 5(1939), pp. 184–206.

[2] The referee has shown that we may retain McShane's definition if we are willing to employ still more general accessory pseudo-extremaloids. I am indebted to the referee for additional suggestions which led to a more elegant formulation.

[3] This is not a vital omission. It may be shown, by a simple modification of a proof of M. Morse (*Calculus of Variations in the Large*, American Mathematical Society Colloquium Publications, vol. 18, New York, 1934, p. 109), that there is a neighborhood of a given extremaloid which can be represented by a single coördinate system.

[4] Right and left limits of a function will be indicated by attaching the symbols + and − to the variable.

[5] For brevity's sake we shorten the phrase "of class C^2 between corners of g and having unique right and left limits at corners of g" to "of class C^2 between corners of g".

(3) If the change of parameter $t = t(T)$ (of class C^2 between corners of g) is made on g, the vector $p(t)$ is replaced by $q(T)$, where

$$q_i(T) = k(T)p_i(t(T))$$

with $k(T)$ non-vanishing and of class C^2 between corners of g.

We shall say that the vector $u^i(t)$ is *p-normal* on a subinterval t_3t_4 of t_1t_2 in case

$$p_i(t)\dot{u}^i(t) = 0 \qquad (t_3 \leqq t \leqq t_4).$$

An accessory pseudo-extremaloid $u^i(t)$, $\tau(t)$ for which $u^i(t)$ is p-normal between corners of g will be called a p-normal accessory pseudo-extremaloid.

Definition of conjugate point. The points $\gamma(t_3)$ and $\gamma(t_4)$ $(t_1 \leqq t_3 < t_4 \leqq t_2)$ of the extremaloid g will be said to be *conjugate* in case there is a p-normal accessory pseudo-extremaloid u^i, τ and constants c, d, e, not all zero with $cd \geqq 0$, such that

(i) $u^i(t) \not\equiv 0$ on t_3t_4,

(ii) $eu^i(t_3) = 0$,

(iii) $e\eta^i(t_4) = c\dot{\gamma}^i(t_4^-) + d\dot{\gamma}^i(t_4^+)$, where $\eta^i = u^i + \tau\dot{\gamma}^i$.

We are now interested in establishing two facts:

(1) Our definition is equivalent to that given by Graves.[6]

(2) For the special case of the non-parametric problem this definition reduces to that of Reid.[7]

To establish (1) suppose that $\gamma(t_4)$ is conjugate to $\gamma(t_3)$ in the sense of Graves. Corresponding to the solution $\eta^i(t)$ of the Jacobi equations (in integral form) we can choose a p-normal accessory pseudo-extremaloid

$$u^i(t) = \eta^i(t) + \rho(t)\dot{\gamma}^i(t), \qquad \tau(t) = -\rho(t)$$

for which $\rho(t_3) = 0$. It is then easily seen that conditions (i), (ii), (iii) hold for u^i, τ. The converse is trivial if $e = 0$. Hence suppose $e \neq 0$. One can choose a function $\rho(t)$ with $\rho(t_3) = 0$ for which $\eta^i = u^i + \rho\dot{\gamma}^i$ is a solution of the Jacobi equations (in integral form), and such that $\rho(x_\theta) = \tau(x_\theta)$ for corner values on t_3t_4. Suppose that $\eta^i \equiv \sigma\dot{\gamma}^i$ on t_3t_4. Then we should have $u^i \equiv (\sigma - \rho)\dot{\gamma}^i$ on t_3t_4. However, since $u^i(t)$ is p-normal and $u^i(t_3) = 0$, it would follow that $u^i(t) \equiv 0$ on t_3t_4, contrary to (i). Thus $\eta^i \not\equiv \sigma\dot{\gamma}^i$ on t_3t_4. It follows readily that $\gamma(t_4)$ is conjugate to $\gamma(t_3)$ in the sense of Graves.

For the non-parametric case we choose, as does McShane, the usual representation of g, i.e., with $y^0 = x = t$, and the vector $p(t)$ to be $(1, 0, \cdots, 0)$. An accessory pseudo-extremaloid $u^\alpha(x)$, $\tau(x)$ is thus p-normal on x_3x_4 if and only if $u^0 =$ constant on x_3x_4. Thus if $\gamma(x_4)$ is conjugate to $\gamma(x_3)$ in our sense, the set $u^i(x)$, $\Omega_{\rho^i}(u, \dot{u})$, $\tau(x_\theta)$ is seen, in view of Lemma 11.1 of McShane, to be a

[6] L. M. Graves, *Discontinuous solutions in space problems of the calculus of variations*, American Journal of Mathematics, vol. 52(1930), p. 18.

[7] W. T. Reid, *Discontinuous solutions in the non-parametric problem of Mayer in the calculus of variations*, American Journal of Mathematics, vol. 57(1935), p. 81.

secondary extremaloid in the sense of Reid. The condition (iii) yields

$$e\eta^0(x_4) = e\tau(x_4) = c + d.$$

It follows that $e \neq 0$. Further use of the condition (iii) gives

$$eu^i(x_4^+) = c[\dot\gamma^i(x_4^-) - \dot\gamma^i(x_4^+)],$$

$$eu^i(x_4^-) = d[\dot\gamma^i(x_4^+) - \dot\gamma^i(x_4^-)].$$

Since $e \neq 0$ we have $cu^i(x_4^-) + du^i(x_4^+) = 0$. If $c + d = 0$, then $c = d = 0 = \tau(x_4)$, and consequently $u^i(x_4^+) = 0$. Thus $\gamma(x_4)$ is conjugate to $\gamma(x_3)$ in the sense of Reid.

Conversely, if $\gamma(x_4)$ is conjugate to $\gamma(x_3)$ in the sense of Reid, the set $(0, u^1(t), \cdots, u^n(t))$, $\tau(t)$ is a p-normal accessory pseudo-extremaloid satisfying (i) and (ii) provided we choose $\tau(x_0) = \tau_0$. We take $\bar e = c + d$, $\bar c = c\tau(x_4)$, $\bar d = d\tau(x_4)$. It is then easily seen, if we use the corner conditions, that the condition (iii) holds. Thus $\gamma(x_4)$ is conjugate to $\gamma(x_3)$ in our sense.

This completes the proofs of (1) and (2).

The use of accessory pseudo-extremaloids does not complicate the *tests* for conjugate points. This is readily seen when one examines the determinants involved. We note also that p-normal accessory pseudo-extremaloids remain "normal" under an admissible transformation of parameter on g.

It seems that this method might be used to simplify some of the details of the theory of discontinuous solutions for more general non-parametric problems of the calculus of variations.

LEHIGH UNIVERSITY.

ON A THEOREM OF P. A. SMITH CONCERNING FIXED POINTS FOR PERIODIC TRANSFORMATIONS

By Samuel Eilenberg

1. **Introduction.** The object of this paper is the discussion and generalization of the following theorem due to Smith:[1]

Let X be a point set in a Cartesian R^m, and Λ a topological transformation of X into itself of a finite and prime[2] period p. If every continuous single-valued image in X of every sphere of dimension $\leqq pm - m - 1$ is deformable in X to a point, then Λ leaves fixed at least one point.

The homotopy condition in this theorem is going to be replaced by a homology condition, expressed in terms of true cycles[3] in X with coefficients from a commutative ring R with a unit element.

Given $x \epsilon R$ and an integer n, we shall say that x is an *inverse* of n if $nx = 1$ (1 being the unit element of R).

Theorem I. *Let X be a metric separable space of finite dimension, and Λ a topological transformation of X into itself of a finite and prime period p. Let R be a commutative ring with a unit element, which does not contain an inverse of p. If every true cycle in X with coefficients in R bounds in X, then Λ leaves fixed at least one point.*

If we take R to be the ring of all integers reduced mod q ($q = 0, 2, 3, \cdots$), it is easy to verify that the prime p has no inverse in R if and only if q is a multiple of p. We therefore obtain

Theorem Ia. *Let X be a metric separable space of finite dimension, and Λ a topological transformation of X into itself of a finite and prime period p. Let q be any multiple of p (including 0). If every true cycle in X with coefficients* mod q *bounds in X, then Λ leaves fixed at least one point.*

True chains and true cycles can be replaced by singular chains and singular cycles throughout. Moreover, if X is a subset of a Cartesian R^m, then only the

Received December 7, 1939.

[1] P. A. Smith, *A theorem on fixed points for periodic transformations*, Annals of Math., vol. 35(1934), pp. 572–578.

[2] It is clear from Smith's proof that p must be a prime though this is not stated in his theorem.

[3] A sequence $C^n = \{c_1^n, c_2^n, \cdots\}$ is called an n-dimensional *true chain* in X if there exist a compact subset Y of X and a sequence of numbers $\epsilon_i \to 0$ such that c_i^n is an n-dimensional ϵ_i-chain in Y. C^n is a *true cycle* if $\partial C^n = 0$, where $\partial C^n = \{\partial c_1^n, \partial c_2^n, \cdots\}$ and ∂ is the usual boundary operator. C^n *bounds* if $C^n = \partial C^{n+1}$ for some true chain C^{n+1} in X. It is convenient in this paper to accept the convention that a true 0-chain $C^0 = \{c_1^0, c_2^0, \cdots\}$ is a 0-cycle only if the sum of coefficients in c_i^0 is 0 for $i = 1, 2, \cdots$.

hypothesis concerning cycles of dimension $\leqq pm - m - 1$ is used in the proofs. We therefore obtain the following statement:

Let X be a point set in a Cartesian R^m, and Λ a topological transformation of X into itself of a finite and prime period p. Let R be a commutative ring with a unit element, which does not contain an inverse of p. If every singular cycle in X of dimension $\leqq pm - m - 1$ and coefficients in R bounds in X, then Λ leaves fixed at least one point.

This generalizes the original theorem of Smith.[4] A similar statement with true cycles instead of singular cycles would be a direct consequence of Theorem I. In fact, $p \geqq 2$ implies $pm - m - 1 \geqq 2m - m - 1 = m - 1$, and true cycles of dimension $\geqq m$ in X bound in X because $X \subset R^m$.

In §2 we introduce chain-transformations which are essential for the whole proof. The algebraic condition concerning R and p is introduced in a lemma in §3. §4 contains the statement and proof of the principal lemma. §5 is a digression: using the just-mentioned lemma we generalize and prove again a theorem of Borsuk[5] concerning antipodes on the n-sphere. The proof of Theorem I is given in §6; it differs from the proof of Smith only in details.

§7 deals with a finite group of homeomorphisms G acting on X, instead of a single periodic transformation. If we use ordinary methods of the theory of finite groups, a generalization of Theorem I is available.

The condition that X should be of finite dimension is justified in §8. We construct a metric separable space (of infinite dimension) in which every compact subset is contractible to a point, and which has a transformation of period 2 without a fixed point.

In §9 we construct examples justifying the algebraic condition in Theorem Ia. In particular, for every prime p we construct a 2-dimensional infinite complex K which has a transformation of period p without a fixed point, and in which every cycle mod q bounds provided q is not a multiple of p. An argument is also given showing that no finite complex can have this property.

2. **Chain transformations.** Let X be a metric space; K, a finite geometrical cell-complex with oriented cells.

Let R be a commutative ring with a unit element. In the space X we shall consider true chains and true cycles[3] with coefficients in R. In the complex K we shall consider ordinary chains and cycles, also with coefficients in R. Since R has a unit element, every oriented n-cell σ^n of K can be considered as an n-chain.

[4] We use the well-known argument that if every continuous single-valued image in X of every sphere of dimension $\leqq k$ is deformable in X into a point, then every continuous single-valued image in X of every polyhedron of dimension $\leqq k$ is deformable in X into a point. The word *image* should be understood as *transformation*, not as *image-set*.

[5] K. Borsuk, Fund. Mathematicae, vol. 20(1933), pp. 177–190, Satz I, p. 178; Alexandroff-Hopf, *Topologie*, I, Berlin, 1935, p. 483.

Suppose that to each n-chain γ^n in K ($n = 0, 1, 2, \cdots$) there corresponds a true n-chain $T(\gamma^n) = \{c_1^n, c_2^n, \cdots\}$ in X such that

(1) $T(\gamma_1^n - \gamma_2^n) = T(\gamma_1^n) - T(\gamma_2^n)$,

(2) $T(a\gamma^n) = aT(\gamma^n)$ if $a \in R$,

(3) $T(\partial\gamma^n) = \partial T(\gamma^n)$,

(4) if $n = 0$ then the sum of coefficients in each c_i^0 is equal to the sum of coefficients in γ^0.

Such a correspondence T will be called a *chain-transformation of K into X*. It follows from (1) and (2) that T is entirely determined by its values $T(\sigma^n)$ for each cell σ^n.

Let p be an arbitrary prime. Suppose that an isomorphic transformation[6] $\Lambda(K) = K$ of period p and a transformation $\Lambda(X) = X$ also of period p are given. A chain-transformation T of K into X will be called *invariant with respect to* Λ if

$$\Lambda[T(\gamma^n)] = T[\Lambda(\gamma^n)] \text{ for every chain } \gamma^n \text{ in } K.$$

LEMMA 1. *Let K' be a closed subcomplex of K invariant under Λ (i.e., such that $\Lambda(K') = K'$) and such that $K - K'$ is r-dimensional and Λ has no fixed points in $K - K'$. If every true cycle of dimension $< r$ in X bounds, then every chain-transformation T of K' into X invariant with respect to Λ can be extended to a chain-transformation of K in X invariant with respect to Λ.*

Proof. Let us denote the identity transformation by Λ^0 and let us put

$$\Lambda^i = \Lambda\Lambda^{i-1} \qquad (i = 1, 2, \cdots).$$

Since the transformation Λ has no fixed point in $K - K'$ and the period p of Λ is a prime, no one of the transformations $\Lambda^1, \Lambda^2, \cdots, \Lambda^{p-1}$ has a fixed point in $K - K'$. It follows that (changing the orientations if necessary) all the k-cells of $K - K'$ can be arranged without repetition in the following matrix ($k = 0, 1, \cdots, r$).

$$\begin{array}{l}
\sigma_1^k = \Lambda^0(\sigma_1^k),\ \Lambda^1(\sigma_1^k),\ \cdots,\ \Lambda^{p-1}(\sigma_1^k),\\
\sigma_2^k = \Lambda^0(\sigma_2^k),\ \Lambda^1(\sigma_2^k),\ \cdots,\ \Lambda^{p-1}(\sigma_2^k),\\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
\sigma_{i_k}^k = \Lambda^0(\sigma_{i_k}^k),\ \Lambda^1(\sigma_{i_k}^k),\ \cdots,\ \Lambda^{p-1}(\sigma_{i_k}^k).
\end{array}$$

Let K^k be the subcomplex of K consisting of all cells of dimension $\leqq k$.

Let x be a point of X and $C^0 = \{x, x, \cdots\}$ the true 0-chain corresponding to it. Taking

$$T[\Lambda^j(\sigma_i^0)] = \Lambda^j(C^0) \qquad (i = 1, 2, \cdots, i_0;\ j = 0, 1, \cdots, p-1),$$

we clearly obtain an extension of T to a chain-transformation of $K' + K^0$ into X invariant with respect to Λ.

Suppose that T has already been extended to a chain-transformation of K'

[6] Isomorphic transformation = homeomorphism carrying cells into cells and linear on each cell.

$+ K^{k-1}$ invariant with respect to Λ $(0 < k \leqq r)$. Since the true $(k-1)$-cycle $T(\partial\sigma_i^k)$ bounds in X, there is a true chain C_i^k such that $\partial C_i^k = T(\partial\sigma_i^k)$ $(i = 1, 2, \cdots, i_k)$. Taking

$$T[\Lambda^j(\sigma_i^k)] = \Lambda^j(C_i^k) \qquad (i = 1, 2, \cdots, i_k; j = 0, 1, \cdots, p-1),$$

we clearly obtain an extension of T to a chain-transformation of $K' + K^k$ into X invariant with respect to Λ. Since $K - K'$ is r-dimensional, we have $K = K' + K^r$. This completes the proof.

The situation becomes simpler when $X = \bar{K}$ and the mappings $\Lambda(X) = X$ and $\Lambda(K) = K$ actually coincide.[7] The simplest example of a chain-transformation of K into $\bar{K}$ invariant with respect to Λ can be defined as follows:

Let K_1 be a simplicial subdivision of K such that Λ remains simplicial with respect to K_1. Let $K_2, K_3, \cdots$ be a sequence of successive barycentric subdivisions of K_1. To each chain γ^n in K and each $i = 1, 2, \cdots$ there corresponds a chain c_i^n in K_i which is the "subdivision" of γ^n. The transformation $T_0(\gamma^n) = \{c_1^n, c_2^n, \cdots\}$ is clearly a chain-transformation of K into $\bar{K}$ invariant with respect to Λ.

3. Simple n-circuits. A finite cell-complex K will be called a *simple n-circuit* (with respect to the commutative ring R with unit element) if

(1) K is n-dimensional;[8]

(2) every k-cycle γ^k in K bounds if $k < n$;

(3) the n-chain $\gamma_0^n = \sum \sigma^n$, where the summation is extended to all (oriented) n-cells of K, is a cycle;

(4) every n-cycle of K is of the form $a\gamma_0^n$, where $a \in R$.

The cycle γ_0^n will be called the *basis cycle* of K.

LEMMA 2. *Let K be a simple n-circuit with respect to the commutative ring R with unit element, and let γ_0^n be the basis cycle of K. Let p be a prime number which has no inverse in the ring R, and Λ an isomorphic transformation of K into itself with the period p with no fixed points. Then, for every chain-transformation T of K into $\bar{K}$ invariant with respect to Λ, the true cycle $T(\gamma_0^n)$ does not bound in $\bar{K}$.*

Proof. Since γ_0^n is a basis cycle, we have

$$\Lambda(\gamma_0^n) = \eta\gamma_0^n,$$

where $\eta = \mp 1$ according to whether Λ changes the orientation of the n-cells or not. Since p is prime and Λ has no fixed point, it follows that none of the mappings $\Lambda^1 = \Lambda, \Lambda^2, \cdots, \Lambda^{p-1}$ has a fixed point. The n-cells of K can therefore be arranged without repetition in the following matrix:

$$\begin{array}{l}
\sigma_1^n = \eta^0\Lambda^0(\sigma_1^n),\ \eta^1\Lambda^1(\sigma_1^n),\ \cdots,\ \eta^{p-1}\Lambda^{p-1}(\sigma_1^n),\\
\sigma_2^n = \eta^0\Lambda^0(\sigma_2^n),\ \eta^1\Lambda^1(\sigma_2^n),\ \cdots,\ \eta^{p-1}\Lambda^{p-1}(\sigma_2^n),\\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
\sigma_{i_n}^n = \eta^0\Lambda^0(\sigma_{i_n}^n),\ \eta^1\Lambda^1(\sigma_{i_n}^n),\ \cdots,\ \eta^{p-1}\Lambda^{p-1}(\sigma_{i_n}^n).
\end{array}$$

[7] K being a complex, $\bar{K}$ denotes the polyhedron for which K is the cell-division.

[8] This condition may be dropped because it is a consequence of (3) and (4).

Taking $\gamma_1^n = \sigma_1^n + \sigma_2^n + \cdots + \sigma_{i_n}^n$, we have

$$\gamma_0^n = \sum_{i=0}^{p-1} \eta^i \Lambda^i(\gamma_1^n).$$

We shall consider the closed interval $E = [0, 1]$ as a complex consisting of two 0-cells 0 and 1 and one 1-cell ϵ oriented so that $\partial\epsilon = 0 - 1$. The Cartesian product $L = K \times E$ will be considered as a complex with cells of the form $\sigma^k \times 0$ (oriented as σ^k), $\sigma^k \times 1$ (oriented as $-\sigma^k$) and $\sigma^k \times \epsilon$ (oriented so as to have $\partial(\sigma^k \times \epsilon) = \partial\sigma^k \times \epsilon + \sigma^k \times \partial\epsilon$). Let L^n be the subcomplex of L consisting of all cells of dimension $\leqq n$.

Given a point $(x, t) \in L^n$ ($x \in K$, $t \in E$), we put $\Lambda(x, t) = (\Lambda(x), t)$. We obtain this way an isomorphic mapping of L^n into itself of period p.

Since $\partial(\gamma_0^n \times \epsilon) = \gamma_0^n \times 0 - \gamma_0^n \times 1$, we have

$$(*) \qquad \sum_{i=1}^{p-1} \eta^i \Lambda^i[\partial(\gamma_1^n \times \epsilon)] = \gamma_0^n \times 0 - \gamma_0^n \times 1.$$

Taking $L' = K \times 0 + K \times 1$ and

$$T_1(\gamma^k \times 0) = T_0(\gamma^k), \qquad T_1(\gamma^k \times 1) = T(\gamma^k),$$

we obtain a chain-transformation T_1 of L' into $\bar{K}$ invariant with respect to Λ. By Lemma 1 we may assume that T_1 has been extended to a chain-transformation of L^n into $\bar{K}$ invariant with respect to Λ. From (*) it therefore follows that

$$\sum_{i=0}^{p-1} \eta^i \Lambda^i[T_1(\partial(\gamma_1^n \times \epsilon))] = T_0(\gamma_0^n) - T(\gamma_0^n).$$

Assuming that $T(\gamma_0^n) \sim 0$ in $\bar{K}$, we get

$$\sum_{i=0}^{p-1} \eta^i \Lambda^i(C^n) \sim T_0(\gamma_0^n),$$

where $C^n = T_1[\partial(\gamma_1^n \times \epsilon)]$ is a true n-cycle in $\bar{K}$. This implies[9] the existence of an n-cycle γ^n in K such that

$$(**) \qquad \sum_{i=0}^{p-1} \eta^i \Lambda^i(\gamma^n) = \gamma_0^n.$$

Since γ_0^n is the basis cycle of K, we have $\gamma^n = a\gamma_0^n$ for some $a \in R$. Since $\Lambda^i(\gamma_0^n) = \eta^i\gamma_0^n$, it follows that $\eta^i\Lambda^i(\gamma^n) = a\gamma_0^n$. Substituting into (**) we get $pa\gamma_0^n = \gamma_0^n$ and finally $pa = 1$. The element $a \in R$ is therefore an inverse for p, and this contradicts our hypothesis.

4. **Principal lemma.** We next prove the following

LEMMA. *Let R be a commutative ring with a unit element and p a prime number which has no inverse in R. Let K be a simple n-circuit (with respect to R) and* Λ

[9] We are using here the argument that ordinary homology in $\bar{K}$ can be replaced by ϵ-homology in $\bar{K}$ for ϵ small enough.

an isomorphic transformation of K into itself with the period p and no fixed points. Let X be a metric space such that every true k-cycle in X with coefficients in R bounds for $k < n$, and let Λ be a transformation of X into itself with the period p. For every continuous transformation $f(X) \subset K$ such that $f\Lambda = \Lambda f$ there is a true n-cycle C^n in X (coef. R) such that $f(C^n)$ does not bound in $\bar{K}$.

Proof. By Lemma 1 there is chain transformation T of K into X which is invariant with respect to Λ. Taking fT we obtain a chain transformation of K into $\bar{K}$ invariant with respect to Λ. By Lemma 2 the true cycle $f[T(\gamma_0^n)]$ does not bound in $\bar{K}$, where γ_0^n is the basis cycle for K.

5. **Borsuk's theorem.** If K is a simple n-circuit with respect to the ring R with a unit element, then every continuous mapping $f(\bar{K}) \subset \bar{K}$ has a uniquely determined degree which is an element of R. This degree is 0 if and only if $f(C^n)$ bounds for every true n-cycle C^n in $\bar{K}$ (coef. R). Using the principal lemma, we therefore obtain the following

THEOREM II. *Let K be a simple n-circuit with respect to the commutative ring R with a unit element. Let p be a prime number which has no inverse in the ring R, and Λ an isomorphic transformation of K into itself with the period p and without a fixed point. Then every continuous transformation $f(\bar{K}) \subset \bar{K}$ such that $f\Lambda = \Lambda f$ has a degree $\neq 0$.*

The theorem can be applied in the case when R is the ring of all integers reduced mod q, where $q = 0, 2, 3, \cdots$ is an arbitrary multiple of p.

If we assume that $px = 0$ for every x in the ring R, then the statements of Lemma 2 and of the principal lemma can be improved in the following way. The assertion in Lemma 2 that $T(\gamma_0^n)$ does not bound in $\bar{K}$ can be replaced by the stronger one that $T_0(\gamma_0^n) - T(\gamma_0^n)$ bounds in $\bar{K}$. In fact, following the proof of Lemma 2 we obtain a true cycle C^n in $\bar{K}$ such that

$$\sum_{i=0}^{p-1} \eta^i \Lambda^i(C^n) = T_0(\gamma_0^n) - T(\gamma_0^n).$$

If $T_0(\gamma_0^n) - T(\gamma_0^n)$ does not bound, there is (see footnote 9) an n-cycle γ^n in K such that

$$(*) \qquad \sum_{i=0}^{p-1} \eta^i \Lambda^i(\gamma^n) \neq 0.$$

Since γ_0^n is the basis cycle of K, we have $\gamma^n = a\gamma_0^n$ for some $a \in R$. Since $\Lambda^i(\gamma_0^n) = \eta^i\gamma_0^n$, it follows that $\eta^i\Lambda^i(\gamma^n) = a\gamma_0^n$. Substituting into (*), we get $pa\gamma_0^n \neq 0$ and therefore $pa \neq 0$.

In the principal lemma the assertion that $f(C^n)$ does not bound can be replaced by the stronger one that $f(C^n) - T_0(\gamma_0^n)$ bounds. This implies the following theorem analogous to Theorem II.

THEOREM III. *Let K be a simple n-circuit with respect to the commutative ring R with a unit element. Let p be a prime number such that $px = 0$ for all*

$x \epsilon R$, and Λ an isomorphic transformation of K into itself with the period p and without a fixed point. Then every continuous transformation $f(\bar{K}) \subset \bar{K}$ such that $f\Lambda = \Lambda f$ has the degree 1.

In particular, if R is the ring of all integers reduced mod p, we have $px = 0$ for every $x \epsilon R$. We therefore obtain

THEOREM IIIa. *Let K be a simple n-circuit* mod p, *where p is a prime, and Λ an isomorphic transformation of K into itself of the period p and without a fixed point. Then every continuous transformation $f(\bar{K}) \subset \bar{K}$ such that $f\Lambda = \Lambda f$ has the degree* 1 (mod p).

This theorem has been established by K. Borsuk (see footnote 5) in the case where K is an n-dimensional spherical manifold S^n, $p = 2$ and Λ is the antipodic transformation of S^n.[10]

6. **Proof of Theorem I.** Since X is a metric separable space of finite dimension, we may admit that X is a subset of a Cartesian m-dimensional space R^m.

Let us consider the space R^{mp} as the product $R^m \times R^m \times \cdots \times R^m$ (p terms). Every point $x \epsilon R^{mp}$ may then be represented in the form

$$x = (x_1, x_2, \cdots, x_p), \qquad x_i \epsilon R^m \qquad (i = 1, 2, \cdots, p).$$

Taking

$$\Lambda(x) = (x_p, x_1, x_2, \cdots, x_{p-1}),$$

we obtain a linear orthogonal transformation of R^{mp} into itself with the period p. The fixed points of this transformation form an m-dimensional hyperspace I^m defined by the condition $x_1 = x_2 = \cdots = x_p$.

Let L^{mp-m} be the $(mp - m)$-dimensional hyperspace orthogonal to I^m and intersecting I^m at the origin. Let $S = S^{mp-m-1}$ be the surface of the unit sphere in L^{mp-m} with the center at the origin. Since $\Lambda(S) = S$ is a linear and orthogonal transformation, we can represent S as a simplicial complex K for which Λ will be simplicial. The $(mp - m - 1)$-simplices of K can be oriented so that K will be a simple $(mp - m - 1)$-circuit (with respect to the given ring R). Let γ_0^{mp-m-1} be the basis cycle of K.

Every point of $R^{mp} - I^m$ can be projected on $L^{mp-m} - I^m$ parallel to I^m and then projected again on S from the origin as a center. Let us denote the resulting transformation of $R^{mp} - I^m$ into S by f. We clearly have

$$f\Lambda = \Lambda f.$$

Given a point $x \epsilon X \subset R^m$, let us consider the point

$$x' = (x, \Lambda(x), \Lambda^2(x), \cdots, \Lambda^{p-1}(x)) \epsilon R^{mp}.$$

[10] For other generalizations, see G. Hirsch, Bulletin de l'Académie Royale de Belgique (Classe des Sciences), vol. 23(1937), pp. 219–225.

Let X' be the subset of R^{mp} thus defined. It is clear that (1) X' is a homeomorphic image of X, (2) $\Lambda(X') = X'$, (3) Λ has a fixed point on X if and only if $X' \cdot I^m \neq 0$.

Let us assume that $X' \subset R^{mp} - I^m$. By the principal lemma there is in X' a true $(mp - m - 1)$-cycle C such that $f(C)$ does not bound in $\bar{K}$. It follows a fortiori that C does not bound in X' contradicting our hypothesis.

7. **Finite groups of transformation.** From Theorem I and general group-theoretical theorems we shall deduce the following

THEOREM IV. *Let X be a metric separable space of finite dimension and Γ a finite group of order $n > 1$ of homeomorphisms acting on X. Let R be a commutative ring with a unit element which does not contain an inverse for n. If every true cycle in X with coefficients in R bounds in X, then at least one transformation of Γ, in addition to the identity transformation, leaves fixed at least one point.*

In fact, since the product of two numbers having an inverse in R also has inverse in R, there is at least one prime divisor p of n which has no inverse in R. The group Γ contains an element Λ of order p.[11] Using Theorem I, we find that the transformation Λ has a fixed point.

Theorem IV can be applied in the case when R is the group of all integers reduced mod q, where q is 0 or any number not relatively prime to n.

A like generalization of Theorem II and Theorem III is also available.

8. **Infinite dimensional example.** Let X be the space of all sequences

$$x = (x_1, x_2, \cdots)$$

of real numbers such that

$$\sum_{i=1}^{\infty} x_i^2 = 1$$

with the distance defined by the formula

$$|x - x'| = \left[\sum_{i=1}^{\infty} (x_i - x_i')^2\right]^{\frac{1}{2}}.$$

As a subset of Hilbert space, X is separable. Taking

$$\Lambda(x) = (-x_1, -x_2, \cdots)$$

for every $x \in X$, we obtain a transformation of period 2 of X into itself without a fixed point.

[11] See R. D. Carmichael, *Introduction to the Theory of Groups of Finite Order*, p. 69, Corollary III.

Let $x^0 = (x_1^0, x_2^0, \cdots)$ be an arbitrary point of X. The function

$$f(x, t) = (f_1(x, t), f_2(x, t), \cdots),$$

where

$$f_i(x, t) = \{(1 - t)x_i - tx_i^0\}\left\{\sum_{i=1}^{\infty} [(1 - t)x_i - tx_i^0]^2\right\}^{-\frac{1}{2}}$$

is defined and continuous for every $x \in X - x^0$ and $0 \leqq t \leqq 1$, thus deforming the set $X - x^0$ into a point. Every proper subset of X can therefore be deformed into a point. In particular, since X is non-compact, every compact subset is contractible to a point.

Another example is furnished by considering the subset of X consisting of all points with only a finite number of coördinates different from 0. This space can be described as a sum $S^1 + S^2 + \cdots$, where S^n is the n-dimensional spherical surface and S^{n-1} is an equator in S^n.

9. **Complexes K_ν.** Let S be the unit circumference, E the closed interval 0, 1. Given an integer $m > 2$, the Möbius strip mod m, M_m, is obtained from the product $S \times E$ by identifying on the circumference $S \times 1$ points corresponding to each other under the rotation of angle $2\pi/m$. The curve $S \times 0$ of M_m will be called the *free curve*; the curve corresponding to $S \times 1$ will be called the *singular curve* (although if $m = 2$, M_2 is the ordinary Möbius strip and this curve has no singularity).[12]

Given any sequence of integers

$$\nu = (m_1, m_2, \cdots), \qquad m_i \geqq 2,$$

we obtain a 2-dimensional infinite polyhedron K_ν by taking the sequence M_{m_1}, $M_{m_2}, \cdots$ and joining M_{m_i} with $M_{m_{i+1}}$ so that the singular curve of M_{m_i} coincides with the free curve of $M_{m_{i+1}}$.

More precisely we can define K_ν as follows. Let

$$b_0 = 1, \qquad b_i = m_1 m_2 \cdots m_i \qquad (i = 1, 2, \cdots).$$

K_ν is obtained from the half-plane (x, y) defined by the condition $y \geqq 0$, as a result of identification of all points (x_1, y_1), (x_2, y_2) such that

$$x_1 \equiv x_2 \left(\text{mod } \frac{2\pi}{b_i}\right), \qquad i \leqq y_1 = y_2 < i + 1 \qquad (i = 1, 2, \cdots).$$

Given a prime p, the translation $T(x, y) = (x + 2\pi/p, y)$ induces a periodic transformation Λ of period p of K_ν. It can be easily seen that Λ has a fixed point if and only if p is a divisor for some m_i. Hence

(1) *Given any prime p which is not a divisor for m_i $(i = 1, 2, \cdots)$, there is a periodic transformation of K_ν of period p and without a fixed point.*

We assume now that K_ν is represented as an infinite simplicial complex. We clearly have that

(2) *All 0- and 2-cycles in K_ν with any coefficient domain bound.*

[12] Cf. Alexandroff-Hopf, p. 270.

To compute the homology classes of 1-cycles we start with the integer coefficient domain. Let C_i be the simple closed curve corresponding to the line $y = i$ and let γ_i be the corresponding 1-cycle ($i = 0, 1, \cdots$).

The section of K_ν defined by the condition $y \leqq i$ can be deformed into C_i and therefore every 1-cycle in this section is homologous to a multiple of γ_i. It follows that $\gamma_0, \gamma_1, \cdots$ can be chosen as a system of generators for the first homology group. We easily find the relations

$$(*) \qquad \gamma_i \sim m_{i+1}\gamma_{i+1} \qquad (i = 0, 1, \cdots).$$

Any relation $k_0\gamma_0 + k_1\gamma_1 + \cdots + k_n\gamma_n \sim 0$ can be reduced applying (*) to a relation $k\gamma_n \sim 0$. A geometrical argument shows then that $k = 0$. It follows that the relations (*) form a complete system of relations. This gives

(3) *The 1-dimensional homology group with integer coefficients of K_ν is isomorphic with the group H_ν given by the generators $a_0, a_1, \cdots$ and relations*

$$a_i = m_{i+1}a_{i+1} \qquad (i = 0, 1, \cdots).$$[13]

If a number q appears in the sequence ν infinitely many times, then every element of H_ν is of the form qa, where $a \in H_\nu$ and the group H_ν reduced mod q gives the null group. Combining this with (2), we obtain

(4) *If the integer q appears in the sequence $\nu = (m_1, m_2, \cdots)$ infinitely many times, then every cycle* mod *q in K_ν bounds.*

If we take the sequence $\nu = (q, q, \cdots)$, it follows from (1) and (4) that

Given any integer $q \geqq 2$, there is a 2-dimensional infinite polyhedron K such that every cycle mod *q in K bounds and for every prime p which is not a divisor for q there is a periodic transformation of K of period p without a fixed point.*

Given a prime p, let us choose $\nu = (m_1, m_2, \cdots)$ so that no multiple of p appears in ν and every non-multiple of p appears in ν infinitely many times. Using (1) and (4), we obtain

Given any prime p, there is a 2-dimensional infinite polyhedron K such that every cycle mod *q in K bounds for every q which is not a multiple of p and there is a periodic transformation of K of period p without a fixed point.*

The above-described phenomena cannot be illustrated on a 1-dimensional polyhedron K. In fact, if all cycles mod q in K bound for some $q \geqq 2$, then K is connected and without a simple closed curve. This implies that all cycles in K with any coefficient domain bound.

The following argument will show that examples of the type given before are impossible among *finite* polyhedra. In fact, assume that all cycles mod q in K bound for some $q \geqq 2$. Since K is finite, this implies[14] that all the Betti numbers of K are 0 (i.e., all the homology groups of K with integer coefficients are finite). This implies by Lefschetz's "Fixed-point-theorem"[15] that every continuous transformation of K into a subset has a fixed point.

UNIVERSITY OF MICHIGAN.

[13] It can be easily seen that H_ν is isomorphic with the subring of the ring of all rational numbers generated by the sequence $m_1^{-1}, m_2^{-1}, \cdots$.

[14] Alexandroff-Hopf, p. 235.

[15] Alexandroff-Hopf, p. 532.

THE DIFFERENCE OF CONSECUTIVE PRIMES

By P. Erdös

Let p_n denote the n-th prime. Backlund [1][1] proved that, for every positive ϵ and infinitely many n, $p_{n+1} - p_n > (2 - \epsilon) \log p_n$. Brauer and Zeitz [2, 10] proved that $2 - \epsilon$ can be replaced by $4 - \epsilon$. Westzynthius [9] proved that for an infinity of n

$$p_{n+1} - p_n > \frac{2 \log p_n \log \log \log p_n}{\log \log \log \log p_n},$$

and this was improved by Ricci [7] to

$$p_{n+1} - p_n > c_1 \log p_n \log \log \log p_n,$$

where, as throughout the paper, the c's denote positive absolute constants. I [4] showed that

$$p_{n+1} - p_n > c_2 \frac{\log p_n \log \log p_n}{(\log \log \log p_n)^2},$$

and lately Rankin [6] proved

$$p_{n+1} - p_n > c_3 \frac{\log p_n \log \log p_n \log \log \log \log p_n}{(\log \log \log p_n)^2}.$$

In the other direction the best known result is that of Ingham [5] which states that for sufficiently large n

$$p_{n+1} - p_n < p_n^{\frac{38}{61}+\epsilon} < p_n^{\frac{5}{8}}.$$

Thus it is known that

$$\limsup_{n \to \infty} \frac{p_{n+1} - p_n}{\log p_n} = \infty.$$

Very much less is known about

$$A = \liminf \frac{p_{n+1} - p_n}{\log p_n}.$$

Hardy and Littlewood proved a few years ago, by using the Riemann hypothesis, that $A \leqq \frac{2}{3}$, and Rankin recently proved, again by using the Riemann hypothesis, that $A \leqq \frac{3}{5}$. In the present paper we are going to prove—without the Riemann hypothesis—that

$$A < 1 - c_4, \qquad \text{for a certain } c_4 > 0.$$

Received December 12, 1939.

[1] Numbers in brackets refer to the bibliography at the end of the paper.

It seems extremely likely that $A = 0$. In fact, a well-known conjecture states that the equation $p_{n+1} - p_n = 2$ has infinitely many solutions (i.e., there are infinitely many prime twins).

We need two lemmas.

LEMMA 1. *The number of solutions of*

$$a = p_i - p_j, \qquad p_j, p_i \leqq n,$$

does not exceed

$$c_5 \prod_{p \mid a} \left(1 + \frac{1}{p}\right) \frac{n}{(\log n)^2}.$$

The proof is well known ([8], p. 670).

LEMMA 2. *Let c_4 be sufficiently small; then*

$$\sum{}' \prod_{p \mid a} \left(1 + \frac{1}{p}\right) < \frac{1}{6c_5} \log n,$$

where the prime indicates that the summation is extended over the a's of the interval

$$(1 - c_4) \log n \leqq a \leqq (1 + c_4) \log n.$$

Proof. We have

$$\sum{}' \prod_{p \mid a} \left(1 + \frac{1}{p}\right) \leqq \sum_{d < (1+c_4) \log n} \frac{1}{d}\left(\frac{2c_4 \log n}{d} + 1\right)$$

$$< c_6 \log n + \sum_{d < (1+c_4) \log n} \frac{1}{d} < \frac{1}{6c_5} \log n$$

for sufficiently small c_4, and the proof is complete.

Now we can prove our theorem. Denote by $p_1, p_2, \cdots, p_x$ the primes of the interval $\frac{1}{2}n, n$. It follows from the prime number theorem that, for sufficiently large n, $x > (\frac{1}{2} - \epsilon)n/\log n$. It suffices to prove that if n is sufficiently large, then for at least one i

$$p_{i+1} - p_i < (1 - c_4) \log n \qquad (i \leqq x - 1).$$

For then we have

$$\liminf_{r \to \infty} \frac{p_{r+1} - p_r}{\log p_r} \leqq \frac{(1 - c_4) \log n}{\log \frac{1}{2}n} \to 1 - c_4.$$

Write

$$b_1 = p_2 - p_1, b_2 = p_3 - p_2, \cdots, b_{x-1} = p_x - p_{x-1}.$$

Evidently

$$\sum_{i=1}^{x-1} b_i \leqq \tfrac{1}{2}n.$$

From Lemmas 1 and 2 it follows that the number of b's lying in the interval

$$(1 - c_4) \log n \leqq b \leqq (1 + c_4) \log n$$

does not exceed

$$c_5 \frac{n}{(\log n)^2} \sum{}' \prod_{p|a} \left(1 + \frac{1}{p}\right) < \frac{n}{6 \log n}.$$

Hence if $b_i < (1 - c_4) \log n$ had no solution, we should obtain

$$\sum_{i=1}^{x-1} b_i > \frac{n}{6 \log n} (1 - c_4) \log n + (\tfrac{1}{3} - \epsilon) \frac{n}{\log n} (1 + c_4) \log n$$

$$= \tfrac{1}{2}n(1 - 2\epsilon) + (\tfrac{1}{6} - \epsilon)c_4 n > \tfrac{1}{2}n.$$

This is an evident contradiction and the theorem is proved.

Denote by $q_1 < q_2 < \cdots < q_y$ the primes not exceeding n. Cramér [3] proved by aid of the Riemann hypothesis that

$$\sum_{i=1}^{y-1} (q_{i+1} - q_i) = O\left(\frac{n}{\log \log n}\right) \quad (q_{i+1} - q_i > (\log q_i)^3).$$

It might be conjectured that the following stronger result also holds:

$$\sum_{i=1}^{y-1} (q_{i+1} - q_i)^2 = O(n \log n).$$

This result if true must be very deep. I could not even prove the following very much more elementary conjecture: Let n be any integer and let $0 < a_1 < a_2 < \cdots < a_x < n$ be the $\varphi(n)$ integers relatively prime to n; then

$$\sum_{i=1}^{x-1} (a_{i+1} - a_i)^2 < c_6 \frac{n^2}{\varphi(n)}.$$

Bibliography

1. R. J. Backlund, *Über die Differenzen zwischen den Zahlen die zu den n ersten Primzahlen teilerfremd sind*, commemoration volume in honor of E. L. Lindelöf, Helsingfors, 1929.
2. A. Brauer and H. Zeitz, *Über eine zahlentheoretische Behauptung von Legendre*, Sitz. Berliner Math. Ges., vol. 29(1930), pp. 116–125.
3. H. Cramér, *On the difference between consecutive primes*, Acta Arithmetica, vol. 2(1937), pp. 23–45.
4. P. Erdös, *On the difference of consecutive primes*, Quarterly Journal of Math., Oxford Series, vol. 6(1935), pp. 124–128.
5. A. E. Ingham, *On the difference between consecutive primes*, Quarterly Journal of Math., Oxford Series, vol. 8(1937), pp. 255–266.
6. R. A. Rankin, *The difference between consecutive primes*, Journal London Math. Soc., vol. 13(1938), pp. 242–247.

7. G. Ricci, *Ricerche aritmetiche sui polinomi*, II: *Intorno a una proposizione non vera di Legendre*, Rend. Circ. Mat. di Palermo, vol. 58(1934).

8. L. Schnirelmann, *Über additive Eigenschaften von Zahlen*, Math. Annalen, vol. 107(1933), pp. 649–690.

9. E. Westzynthius, *Über die Verteilung der Zahlen die zu den n ersten Primzahlen teilerfremd sind*, Comm. Phys. Math. Soc. Sci. Fenn., Helsingfors, vol. 5(1931), no. 25, pp. 1–37.

10. H. Zeitz, *Elementare Betrachtung über eine zahlentheoretische Behauptung von Legendre*, Berlin, 1930.

Institute for Advanced Study.

A CLASS OF INEQUALITIES

By Willis B. Caton

Chapter I

F. Carlson[1] has proved the following inequality:

$$\left(\sum_1^\infty a_n\right)^4 < \pi^2 \sum_1^\infty a_n^2 \cdot \sum_1^\infty n^2 a_n^2, \qquad a_n \geqq 0. \tag{1.1}$$

He has also shown that π^2 is the best constant. Finally, he has shown that the inequality (1.1) may be considered as a limiting case of a Hölder inequality. This last statement is of particular importance to us and we shall return to it shortly. It is known that Carlson discovered (1.1) while engaged in studies in the theory of analytic functions. We can show that (1.1) has interesting consequences in this theory. For instance, consider the function

$$f(z) = \sum_1^\infty a_n z^n$$

with radius of convergence R. Then, for $r < R$,

$$\{M(r)\}^2 \leqq \pi r \mathfrak{M}_2(f)\mathfrak{M}_2(f'),$$

where $M(r)$ is the maximum value of $|f|$ on $|z| = r$ and $\mathfrak{M}_2(f)$, $\mathfrak{M}_2(f')$ represent the quadratic means of f and f' respectively on the circle. To obtain this result we need only take $c_n = |a_n| r^n$ and use (1.1).

We shall now show how the Carlson result may be considered as a limiting case of a Hölder inequality. The Hölder inequality gives

$$\sum_1^\infty a_n \leqq \left(\sum_1^\infty a_n^2\right)^{\frac14}\left(\sum_1^\infty n^{2h} a_n^2\right)^{\frac14}\left(\sum_1^\infty n^{-h}\right)^{\frac12}.$$

Putting

$$K(h) = \left(\sum_1^\infty n^{-h}\right)^{\frac12},$$

we can write

$$\sum_1^\infty a_n \leqq K(h)\left(\sum_1^\infty a_n^2\right)^{\frac14}\left(\sum_1^\infty n^{2h} a_n^2\right)^{\frac14},$$

Received December 15, 1939. This paper was a dissertation presented for the degree of Doctor of Philosophy in Yale University and was written under the direction of Professor E. Hille.

[1] F. Carlson, *Une inégalité*, Arkiv för Matematik, Astronomi och Fysik, vol. 25B(1934), no. 1, pp. 1–3.

and this is evidently valid if $h > 1$. If $h \to 1$, $K(h) \to +\infty$, so that we cannot establish the existence of a finite K in this manner with the property that

$$\sum_1^\infty a_n \leqq K\left(\sum_1^\infty a_n^2\right)^{\frac{1}{4}}\left(\sum_1^\infty n^2 a_n^2\right)^{\frac{1}{4}}, \qquad a_n \geqq 0.$$

We have seen in the preceding paragraph how the Carlson inequality may be considered as a limiting case of a Hölder inequality. It is just such limiting cases of the Hölder inequality that we are going to study in this paper.

Let

$$\infty > p_i > 0, \qquad \frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} = 1,$$

$$\beta_1, \beta_2 \geqq 0, \qquad \beta_1 + \beta_2 = \beta_3,$$

$$\alpha_1, \alpha_2 \geqq 0, \qquad \alpha_1 + \alpha_2 = 1.$$

Then, the Hölder inequality gives

$$\sum_1^\infty a_n \leqq \left\{\sum_1^\infty (n^{\beta_1} a_n^{\alpha_1})^{p_1}\right\}^{1/p_1} \left\{\sum_1^\infty (n^{\beta_2} a_n^{\alpha_2})^{p_2}\right\}^{1/p_2} \left\{\sum_1^\infty n^{-\beta_3 p_3}\right\}^{1/p_3}.$$

If we put

$$K(\beta_3, p_3) = \left\{\sum_1^\infty n^{-\beta_3 p_3}\right\}^{1/p_3},$$

we get

$$\sum_1^\infty a_n \leqq K(\beta_3, p_3)\left\{\sum_1^\infty (n^{\beta_1} a_n^{\alpha_1})^{p_1}\right\}^{1/p_1} \left\{\sum_1^\infty (n^{\beta_2} a_n^{\alpha_2})^{p_2}\right\}^{1/p_2},$$

a result valid if $\beta_3 p_3 > 1$. If $\beta_3 p_3 \to 1$, $K(\beta_3, p_3) \to +\infty$.

Then, the main problem of this paper is the consideration of the case where $\beta_3 p_3 = 1$. That is, we want to consider the question of the existence of a $K < \infty$ such that

$$\sum_1^\infty a_n \leqq K\left\{\sum_1^\infty (n^{\beta_1} a_n^{\alpha_1})^{p_1}\right\}^{1/p_1} \left\{\sum_1^\infty (n^{\beta_2} a_n^{\alpha_2})^{p_2}\right\}^{1/p_2} \tag{1.2}$$

for any sequence $\{a_n\}$ with $a_n \geqq 0$ and not all $a_n = 0$. It is evident that in general K will be a function of the parameters. Let us set

$$\alpha_i p_i = \sigma_i \qquad (i = 1, 2),$$

$$\beta_i p_i = \tau_i \qquad (i = 1, 2).$$

We shall refer to the class of inequalities (1.2) as the class H_1 if $\sigma_1 \neq \sigma_2$. In case $\sigma_1 = \sigma_2$ we shall denote this class by H_0.

The following result is evident:

LEMMA 1.1. *The Carlson inequality is of the type H_0 with $p_1 = p_2 = 4$, $p_3 = 2$, $\alpha_1 = \alpha_2 = \frac{1}{2}$, $\beta_1 = \frac{1}{2}$, $\beta_2 = 0$.*

We shall now formulate the main result of this chapter.

THEOREM 1.1. *If an inequality of class H_0 exists and $\tau_1 > \tau_2$, then*

$$K \geqq \Gamma\left(\frac{\sigma - 1 - \tau_2}{(\sigma - 1)(\tau_1 - \tau_2)}\right)^{1/p_3} \Gamma\left(\frac{\tau_1 + 1 - \sigma}{(\sigma - 1)(\tau_1 - \tau_2)}\right)^{1/p_3} \Gamma\left(\frac{1}{\sigma - 1}\right)^{-1/p_3}$$

$$\cdot (\tau_1 - \tau_2)^{1-2/p_3}(\sigma - 1 - \tau_2)^{-1/p_1}(\tau_1 + 1 - \sigma)^{-1/p_2}.$$

In order to prove this result we shall first prove a sequence of lemmas. The finite form of the Hölder inequality gives

$$\sum_1^N a_n \left\{\sum_1^N n^{\tau_1} a_n^{\sigma_1}\right\}^{-1/p_1} \left\{\sum_1^N n^{\tau_2} a_n^{\sigma_2}\right\}^{-1/p_2} \leqq \left\{\sum_1^N \frac{1}{n}\right\}^{1/p_3}. \tag{1.3}$$

Now let M_N be the l.u.b. of the left member of (1.3), where $a_n \geqq 0$ ($n = 1, 2, \cdots, N$) and not all the a's are zero. Then we have

LEMMA 1.2. *M_N is finite for every N.*

Let us consider inequalities of the class H_0. Let it be supposed that $\tau_1 = \tau_2 = \tau$ also. Then, we can verify that $\sigma = \tau + 1$.

We consider (1.3) with this choice of parameters and

$$a_n = \frac{1}{n} \qquad (n = 1, 2, \cdots, N).$$

By a simple computation, the left member of (1.3) is

$$\left(\sum_1^N \frac{1}{n}\right)^{1/p_3} \to +\infty, \qquad N \to +\infty.$$

Hence we obtain

LEMMA 1.3. *If an inequality of class H_0 exists, then $\tau_1 \neq \tau_2$.*

Next, we can obtain the following lemma from the fundamental relations.

LEMMA 1.4. *The parameters of H_0 satisfy the following relations*:

(1) $$\sigma\left(\frac{1}{p_1} + \frac{1}{p_2}\right) = 1,$$

(2) $$\frac{1 + \tau_1 - \sigma}{p_1} = \frac{\sigma - \tau_2 - 1}{p_2},$$

(3) $$\sigma > 1,$$

(4) $$\tau_1 > \tau_2 \textit{ implies } \tau_1 > \sigma - 1 > \tau_2,$$

(5) $$1 + \tau_i - \sigma \neq 0 \textit{ (when } H_0 \textit{ exists).}$$

Let us denote the function in the left member of inequality (1.3) by

$$Q(a; N).$$

Then, this function is defined in the domain $a_n \geqq 0$ $(n = 1, 2, \cdots, N)$, the point $a_1 = a_2 = \cdots = a_N = 0$ being excluded. By (1.3) we have

$$M_N = \text{l.u.b. } Q(a; N) < \infty, \qquad M_{N-1} = \text{l.u.b. } Q(a; N-1) < \infty,$$

and evidently also

LEMMA 1.5. $M_{N-1} \leqq M_N$.

Next, one sees easily that we have

LEMMA 1.6. $Q(\rho a; N) = Q(a; N)$, $\rho > 0$.

Thus, all values of $Q(a; N)$ that are assumed in $a_n \geqq 0$ $(n = 1, \cdots, N)$ (not all the a's are zero) are also assumed in

$$a_n \geqq 0, \qquad \sum_1^N a_n = 1.$$

This latter set is a bounded closed set, and the function is continuous over it, so that there exists at least one point $\{a^*\}$ either in the interior or on the boundary and such that

$$Q(a^*; N) = M_N.$$

For finding extreme values taken on in the interior we may use the calculus, but of course, this method will not apply if the extreme value is assumed at a boundary point. Let us consider this second possibility first. To be definite, suppose just one $a_k = 0$ and that $k < N$. Then, a simple rearrangement argument shows that this is not possible.

In case the extreme value is assumed at a boundary point with $a_N = 0$, we have

$$M_N = Q(a; N) = Q(a; N-1) \leqq M_{N-1},$$

and so $M_N \leqq M_{N-1}$. This with Lemma 1.5 implies $M_N = M_{N-1}$. Thus, values of N for which the extreme value is assumed at a boundary point with $a_N = 0$ may be dropped. Finally, the above arguments are valid if any number of the a's less than N is zero.

Since we have shown that we can limit ourselves to interior extremes, we may apply the method of Lagrange. For inequalities of class H_0 we consider

$$Q(a; N) = \left\{\sum_1^N n^{\tau_1} a_n^{\sigma}\right\}^{-1/p_1} \left\{\sum_1^N n^{\tau_2} a_n^{\sigma}\right\}^{-1/p_2}$$

for $a_n \geqq 0$ $(n = 1, 2, \cdots, N)$ and $\sum_1^N a_n = 1$. We want to maximize $Q(a; N)$, that is, to minimize

$$\left\{\sum_1^N n^{\tau_1} a_n^{s}\right\}^{1/p_1} \left\{\sum_1^N n^{\tau_2} a_n^{s}\right\}^{1/p_2}.$$

Let us suppose that

$$\sum_1^N n^{\tau_2} a_n^\sigma = R,$$

where R has a fixed value, and minimize

$$\left\{\sum_1^N n^{\tau_1} a_n^\sigma\right\}^{1/p_1}, \quad \text{i.e.,} \quad \sum_1^N n^{\tau_1} a_n^\sigma.$$

The method of Lagrange leads to the equations

$$\sigma(k^{\tau_1} + \mu k^{\tau_2}) a_k^{\sigma-1} - \lambda = 0,$$

for $k = 1, 2, \cdots, N$, where μ and λ are parameters. Since we have already seen that $\sigma > 1$, we get

LEMMA 1.8. *The a's giving extreme values are of the form*

$$a_n^{\sigma-1} = \frac{\lambda'}{n^{\tau_1} + \mu n^{\tau_2}} \qquad (n = 1, 2, \cdots, N).$$

When $\tau_1 > \tau_2$,

if $\lambda' > 0, \quad \mu > -1,$

if $\lambda' < 0, \quad \mu < -N^{\tau_1 - \tau_2}.$

Our preceding studies lead us to a consideration of the infinite sequence

$$a_n = \frac{1}{(n^{\tau_1} + \mu n^{\tau_2})^{1/(\sigma-1)}} \qquad (n = 1, 2, 3, \cdots)$$

with $\mu > -1$. We note that when an inequality of type H_0 exists, then one must have in particular

LEMMA 1.9.

$$\underset{-1<\mu<\infty}{\text{l.u.b.}}\ Q(\mu) < \infty,$$

where

$$Q(\mu) = S_0(\mu)\{S_1(\mu)\}^{-1/p_1}\{S_2(\mu)\}^{-1/p_2},$$

$$S_0(\mu) = \sum_1^\infty (n^{\tau_1} + \mu n^{\tau_2})^{-1/(\sigma-1)},$$

$$S_1(\mu) = \sum_1^\infty n^{\tau_1}(n^{\tau_1} + \mu n^{\tau_2})^{-\sigma/(\sigma-1)},$$

$$S_2(\mu) = \sum_1^\infty n^{\tau_2}(n^{\tau_1} + \mu n^{\tau_2})^{-\sigma/(\sigma-1)}.$$

Let us consider the convergence of these three series. We shall first investigate the case with $\tau_1 > \tau_2$. We recall that this implies $\tau_1 > \sigma - 1 > \tau_2$.

We begin with the series $S_0(\mu)$. This series converges or diverges with the integral

$$\int_a^\infty (x^{\tau_1} + \mu x^{\tau_2})^{-1/(\sigma-1)}\, dx, \qquad \infty > \mu > -1.$$

By the usual methods of the calculus this integral has a sense so that $S_0(\mu)$ is convergent for $\infty > \mu > -1$.

Next, consider $S_0(\mu)$ in the closed interval

$$-1 + \epsilon \leqq \mu \leqq \frac{1}{\epsilon}, \qquad \epsilon > 0.$$

We may show by the Weierstrass Theorem that $S_0(\mu)$ is uniformly convergent in any such interval.

Similar argument shows that $S_1(\mu)$ and $S_2(\mu)$ are both convergent in $-1 < \mu < \infty$, and uniformly convergent in any $[-1 + \epsilon \leqq \mu \leqq \epsilon^{-1}]$. Hence all three series represent continuous functions in any such interval. Furthermore, the ratio $Q(\mu)$ is clearly continuous and consequently bounded in any such interval. We must investigate the behavior of $Q(\mu)$ as $\mu \to -1+$ and as $\mu \to +\infty$. Elementary considerations show that

$$\lim_{\mu \to -1+} Q(\mu) = 1.$$

The behavior of $Q(\mu)$ for large $\mu > 0$ is not so simply obtained. We shall proceed to prove a sequence of lemmas which will lead to the solution of this problem.

Let us put

$$\varphi(x; \mu) = (x^{\tau_1} + \mu x^{\tau_2})^{-1/(\sigma-1)}.$$

By the method of MacLaurin

$$S_0(\mu) > \int_1^\infty \varphi(x; \mu)\, dx > S_0(\mu) - \varphi(1; \mu),$$

$$\varphi(1; \mu) = O(\mu^{-1/(\sigma-1)}), \qquad \mu \to +\infty.$$

Hence, if the integral appearing above is denoted by $I_0(\mu)$, we get

LEMMA 1.10.

$$S_0(\mu) = I_0(\mu) + O(\mu^{-1/(\sigma-1)}), \qquad \mu \to +\infty.$$

Next, after a suitable transformation of the variable, we have

$$I_0(\mu) = \alpha_0' \mu^{1/\alpha_0'\tau - \nu} \int_\eta^\infty \frac{dy}{(y^{\alpha_0'\tau} + 1)^\nu},$$

with

$$\eta = \mu^{-(\sigma-1-\tau_2)/[(\sigma-1)(\tau_1-\tau_2)]} \to 0, \qquad \mu \to +\infty,$$

$$\alpha_0' = \frac{\sigma - 1}{\sigma - 1 - \tau_2}, \qquad \tau = \tau_1 - \tau_2, \qquad \nu = \frac{1}{\sigma - 1}.$$

Let us put

$$\bar{I}_0(\mu) = \int_\eta^\infty \frac{dy}{(y^{\alpha_0'\tau} + 1)^\nu} \to \int_0^\infty \frac{dy}{(y^{\alpha_0'\tau} + 1)^\nu}, \qquad \mu \to +\infty,$$

and this limit can be expressed in terms of gamma functions by well-known formulas. We find that as $\mu \to +\infty$,

$$(1.4) \quad \bar{I}_0(\mu) \to \frac{\sigma - 1 - \tau_2}{(\sigma - 1)(\tau_1 - \tau_2)} \frac{\Gamma\left(\frac{\sigma - 1 - \tau_2}{(\sigma - 1)(\tau_1 - \tau_2)}\right)\Gamma\left(\frac{\tau_1 + 1 - \sigma}{(\sigma - 1)(\tau_1 - \tau_2)}\right)}{\Gamma\left(\frac{1}{\sigma - 1}\right)}.$$

We may sum up these results as follows:

LEMMA 1.11.

$$I_0(\mu) = \frac{\sigma - 1}{\sigma - \tau_2 - 1}\, \mu^{e_0} \bar{I}_0(\mu),$$

where

$$e_0 = \frac{\sigma - 1 - \tau_1}{(\sigma - 1)(\tau_1 - \tau_2)},$$

and $\bar{I}_0(\infty)$ *is given by* (1.4).

Let us now turn to the series $S_1(\mu)$. We put

$$\varphi_1(x; \mu) = x^{\tau_1}(x^{\tau_1} + \mu x^{\tau_2})^{-\sigma/(\sigma-1)}.$$

When

$$\frac{\sigma(\tau_1 - \tau_2)}{\tau_1} - 1 \leqq 0,$$

$\varphi_1(x; \mu)$ is a decreasing function of x in $(1, \infty)$ for every μ satisfying $\infty > \mu > -1$ and the MacLaurin method gives

$$S_1(\mu) = I_1(\mu) + O(\mu^{-\sigma/(\sigma-1)}), \qquad \mu \to +\infty,$$

where

$$I_1(\mu) = \int_1^\infty x^{\tau_1}(x^{\tau_1} + \mu x^{\tau_2})^{-\sigma/(\sigma-1)}\, dx.$$

When

$$\frac{\sigma(\tau_1 - \tau_2)}{\tau_1} - 1 > 0,$$

we make use of the Euler summation formula and find that

$$S_1(\mu) = I_1(\mu) + O(\mu^{-\tau_1/[(\sigma-1)(\tau_1-\tau_2)]}).$$

Hence we have

LEMMA 1.12.

$$S_1(\mu) = I_1(\mu) + O(\mu^{-\sigma/(\sigma-1)}) \quad or \quad O(\mu^{-\tau_1/[(\sigma-1)(\tau_1-\tau_2)]})$$

according as $\sigma(\tau_1 - \tau_2) - \tau_1 \leqq 0$ *or* > 0.

We apply the same methods to $S_2(\mu)$. We put

$$\varphi_2(x;\mu) = x^{\tau_2}(x^{\tau_1} + \mu x^{\tau_2})^{-\sigma/(\sigma-1)},$$

and

$$I_2(\mu) = \int_1^\infty x^{\tau_2}(x^{\tau_1} + \mu x^{\tau_2})^{-\sigma/(\sigma-1)}\,dx.$$

Next, $\partial\varphi_2/\partial x = 0$ has no roots in $1 < x < \infty$ so that the MacLaurin method gives

LEMMA 1.13.

$$S_2(\mu) = I_2(\mu) + O(\mu^{-\sigma/(\sigma-1)}).$$

We next consider $I_1(\mu)$ and $I_2(\mu)$. Computations of the same type that were used to prove Lemma 1.11 give the following results:

LEMMA 1.14.

$$I_1(\mu) = \frac{\sigma - 1}{\sigma(\tau_1 - \tau_2 + 1) - (\tau_1 + 1)}\,\mu^{e_1}\bar{I}_1(\mu),$$

$$e_1 = \frac{\sigma - 1 - \tau_1}{(\sigma - 1)(\tau_1 - \tau_2)},$$

$$\bar{I}_1(\mu) \to \frac{\sigma(\tau_1 - \tau_2 + 1) - (\tau_1 + 1)}{(\sigma - 1)(\tau_1 - \tau_2)}\,\Gamma\left(\frac{\sigma(\tau_1 - \tau_2 + 1) - (\tau_1 + 1)}{(\sigma - 1)(\tau_1 - \tau_2)}\right)$$

$$\cdot\,\Gamma\left(\frac{\tau_1 + 1 - \sigma}{(\sigma - 1)(\tau_1 - \tau_2)}\right)\Big/\Gamma\left(\frac{\sigma}{\sigma - 1}\right), \qquad \mu \to +\infty.$$

LEMMA 1.15.

$$I_2(\mu) = \frac{\sigma - 1}{\sigma - 1 - \tau_2}\,\mu^{e_2}\bar{I}_2(\mu),$$

$$e_2 = \frac{\sigma(1 - \tau_1 + \tau_2) - (1 + \tau_2)}{(\sigma - 1)(\tau_1 - \tau_2)},$$

$$\bar{I}_2(\mu) \to \frac{\sigma - 1 - \tau_2}{(\sigma - 1)(\tau_1 - \tau_2)}\,\Gamma\left(\frac{\sigma - 1 - \tau_2}{(\sigma - 1)(\tau_1 - \tau_2)}\right)$$

$$\cdot\,\Gamma\left(\frac{\sigma(\tau_1 - \tau_2 - 1) + (\tau_2 + 1)}{(\sigma - 1)(\tau_1 - \tau_2)}\right)\Big/\Gamma\left(\frac{\sigma}{\sigma - 1}\right), \qquad \mu \to +\infty.$$

Now, by Lemmas 1.10 to 1.15 inclusive, we have with obvious notation

$$S_i(\mu) = \alpha_i' \mu^{e_i} \bar{I}_i(\mu) + \eta_i(\mu) \qquad (i = 0, 1, 2).$$

But

$$\begin{aligned} Q(\mu) &= S_0(\mu)\{S_1(\mu)\}^{-1/p_1}\{S_2(\mu)\}^{-1/p_2} \\ &= \mu^{e_0 - e_1/p_1 - e_2/p_2} \frac{\alpha_0' \bar{I}_0(\mu) + \omega_0(\mu)}{\{\alpha_1' \bar{I}_1(\mu) + \omega_1(\mu)\}^{1/p_1}\{\alpha_2' \bar{I}_2(\mu) + \omega_2(\mu)\}^{1/p_2}}, \end{aligned}$$

and we may verify that

$$\omega_i(\mu) \equiv \frac{\eta_i(\mu)}{\mu^{e_i}} \to 0, \qquad \mu \to +\infty \qquad (i = 0, 1, 2).$$

Also, we show that

$$e_0 - \frac{e_1}{p_1} - \frac{e_2}{p_2} \equiv 0.$$

Hence, for

$$Q(\infty) = \lim_{\mu \to +\infty} Q(\mu),$$

we get

$$\alpha_0' \bar{I}_0(\infty)\{\alpha_1' \bar{I}_1(\infty)\}^{-1/p_1}\{\alpha_2' \bar{I}_2(\infty)\}^{-1/p_2}.$$

Next, by Lemmas 1.11, 1.14, and 1.15 we have

LEMMA 1.16. *As* $\mu \to +\infty$,

$$\begin{aligned} Q(\mu) \to \Gamma\left(\frac{\sigma - 1 - \tau_2}{(\sigma - 1)(\tau_1 - \tau_2)}\right)^{1/p_3} \Gamma\left(\frac{\tau_1 + 1 - \sigma}{(\sigma - 1)(\tau_1 - \tau_2)}\right)^{1/p_3} \Gamma\left(\frac{1}{\sigma - 1}\right)^{-1/p_3} \\ \cdot (\tau_1 - \tau_2)^{1 - 2/p_3}(\sigma - 1 - \tau_2)^{-1/p_1}(\tau_1 + 1 - \sigma)^{-1/p_2}. \end{aligned}$$

We are now in a position to prove the fundamental Theorem 1.1. Let us assume we have an inequality of class H_0 with $\tau_1 > \tau_2$ and

$$K < Q(\infty).$$

This is clearly not possible since by choosing a sufficiently large μ the series

$$\sum_1^\infty (n^{\tau_1} + \mu n^{\tau_2})^{-1/(\sigma - 1)}$$

would have a $Q(\mu)$ as near to $Q(\infty)$ as we wish and hence a $Q(\mu) > K$ for sufficiently large μ, but this is a contradiction and Theorem 1.1 is established.

Chapter II

By use of methods similar to those in Chapter I, but differing in detail, Gabriel[2] has established the inequality

$$\left(\sum_1^\infty a_n\right)^{2p} \leqq A(p, \delta) \sum_1^\infty n^{p-1+\delta} a_n^p \sum_1^\infty n^{p-1-\delta} a_n^p, \tag{2.1}$$

[2] R. M. Gabriel, *An extension of an inequality due to Carlson*, Journal of the London Math. Soc., vol. 12(1937), pp. 130–132.

for a series $a_n \geqq 0$ and not all $a_n = 0$. The above result is valid if $p > 1$, $0 < \delta \leqq p - 1$. In this paper he gives an expression for $A(p, \delta)$ and observes that (2.1) is a generalization of Carlson's result and that (2.1) gives Carlson's result with the correct constant ($p = 2$, $\delta = 1$).

The Gabriel result thus crosses our own class H_0 and we shall see later how they are related. We shall apply Gabriel's methods and establish the existence of H_0. Our first main result is the following:

THEOREM 2.1. *If $\tau_1 > \tau_2$, then H_0 exists and*

$$K \leqq \left(\frac{p_2}{p_1}\right)^{1/p_2}\left(1 + \frac{p_1}{p_2}\right)^{1/\sigma}(\tau_1 - \tau_2)^{-1/p_3}\tilde{I}^{1/p_3},$$

$$\tilde{I} = \Gamma\left(\frac{\sigma - 1 - \tau_2}{(\sigma - 1)(\tau_1 - \tau_2)}\right)\Gamma\left(\frac{\tau_1 + 1 - \sigma}{(\sigma - 1)(\tau_1 - \tau_2)}\right)\Big/\Gamma\left(\frac{1}{\sigma - 1}\right).$$

We now proceed with the proof of this theorem. Let us put

$$S_0 = \sum_1^N a_n, \quad \text{fixed},$$

$$S_1 = \sum_1^N n^{\tau_1} a_n^\sigma, \qquad S_2 = \sum_1^N n^{\tau_2} a_n^\sigma,$$

and we want to study the ratio

$$S_0 S_1^{-1/p_1} S_2^{-1/p_2}. \tag{2.2}$$

Let us suppose we have a set $(a_1, \cdots, a_N)$ for which the ratio takes on an extreme value with $\sum_1^N a_n = S_0$ fixed. By results of Chapter I we may suppose that no $a_k = 0$. Hence if $x > 0$ and $x < \min a_n$ $(n = 1, 2, \cdots, N)$, then

$$a_1, \cdots, a_m + x, \cdots, a_n - x, \cdots, a_N$$

are all positive and the sum is S_0. We then form the function (2.2) of the above set and it will have an extreme value when and only when $S_1^{-1/p_1} S_2^{-1/p_2}$ has, S_0 being independent of x. Hence

$$\frac{\partial}{\partial x}[S_1^{-1/p_1} S_2^{-1/p_2}]_{x=0} = 0,$$

and this implies that

$$a_n = k(\bar{\alpha} n^{\tau_2} + n^{\tau_1})^{-1/(\sigma-1)},$$

where $n = 1, 2, \cdots, N$, k a positive parameter, and $\bar{\alpha} = p_1 S_2/(p_2 S_1) > 0$.

We now compute

$$S_0 S_1^{-1/p_1} S_2^{-1/p_2} = \mathfrak{A}\left\{\sum_1^N (\bar{\alpha} n^{\tau_2} + n^{\tau_1})^{-1/(\sigma-1)}\right\}^{(\sigma-1)/\sigma},$$

where

$$\mathfrak{A} = \left(\frac{\bar{\alpha} p_2}{p_1}\right)^{1/p_2}\left(1 + \frac{p_1}{p_2}\right)^{1/\sigma}.$$

Hence

$$S_0 S_1^{-1/p_1} S_2^{-1/p_2} \leqq \mathfrak{A}\left\{\int_0^\infty (\bar{\alpha}x^{\tau_2} + x^{\tau_1})^{-1/(\sigma-1)}\, dx\right\}^{(\sigma-1)/\sigma}.$$

If the integral is computed by methods used in Chapter I, the right member of this inequality is found to equal

$$\left(\frac{p_2}{p_1}\right)^{1/p_2}\left(1+\frac{p_1}{p_2}\right)^{1/\sigma} \bar{\alpha}^{e}(\tau_1-\tau_2)^{-(\sigma-1)/\sigma}\,\bar{I}^{(\sigma-1)/\sigma},$$

where

$$e \equiv \frac{\sigma-1-\tau_1}{\sigma(\tau_1-\tau_2)} + \frac{1}{p_2} \equiv 0.$$

Since $(\sigma - 1)/\sigma = 1/p_3$, we have for every N

$$S_0 S_1^{-1/p_1} S_2^{-1/p_2} \leqq \left(\frac{p_2}{p_1}\right)^{1/p_2}\left(1+\frac{p_1}{p_2}\right)^{1/\sigma}(\tau_1-\tau_2)^{-1/p_3}\,\bar{I}^{1/p_3}.$$

The proof of Theorem 2.1 is complete since the right member of this inequality is independent of N.

Let us denote the constant in Theorem 1.2 by K_1 and that in Theorem 2.1 by K_2. By these theorems, if $\tau_1 > \tau_2$ our inequality of class H_0 exists and $K_1 \leqq K \leqq K_2$, where K is the best constant. Simple algebra shows that $K_1 \equiv K_2$, so that we have

THEOREM 2.2. *If $\tau_1 > \tau_2$, the best constant in H_0 is*

$$\left(\frac{p_2}{p_1}\right)^{1/p_2}\left(1+\frac{p_1}{p_2}\right)^{1/\sigma}(\tau_1-\tau_2)^{-1/p_3}\,\bar{I}^{1/p_3}.$$

The results of Chapters I and II are true if $\tau_2 > \tau_1$: we need only interchange the subscripts in the theorems.

Finally, the first Gabriel inequality (2.1) may be written as one of class H_0 with $\tau_1 > \tau_2$, $\alpha_1 = \alpha_2 = \frac{1}{2}$, $p_1 = p_2 = 2\sigma$.

The second Gabriel inequality[3] with

$$\frac{p-1}{p_1} - \delta \geqq 0$$

may be written as an inequality of class H_0 with

$$\tau_1 > \tau_2, \qquad p_1 \geqq p_2.$$

[3] Op. cit.

We shall just mention a result that can be obtained from a theorem of Hardy and Littlewood,[4] namely:

THEOREM 2.3. *If, in addition to the usual restrictions, the parameters of* H_1 *(or* H_0*) satisfy*

$$p_1 > 2, \qquad p_2 > 2, \qquad \frac{1}{p_1} + \frac{1}{p_2} \geqq \frac{1}{2}; \qquad \alpha_1 = \alpha_2 = \frac{1}{2},$$

$$2\left(\beta_1 + \frac{1}{p_1}\right) < 1, \qquad 0 < 2\left(\beta_2 + \frac{1}{p_2}\right) - 1 < 1,$$

then H_1 *(or* H_0*) exists with a finite* K *which is at most*

$$\left\{\frac{2\pi}{\sin 2\pi\,(\beta_1 + 1/p_1)}\right\}^{\frac{1}{2}}.$$

Chapter III

In Chapter II we were able to give a general existence theorem for inequalities of the class H_0. We shall now do the same thing for inequalities of class H_1. The main result of this chapter is the following

THEOREM 3.1. *If* $\sigma_1 - 1 > \sigma_2 - 1 > 0$ *and* $\tau_1 \gtrless \sigma_1 - 1$, *then* H_1 *exists with a finite* K. *In case* $\tau_1 = \sigma_1 - 1$, *there is no inequality.*

Our proof of this result is rather long, and so we shall break it up into a sequence of lemmas. We might add that the argument in this chapter will be similar to that in Chapter II, but in addition, we shall find it appropriate to make use of certain results of Mellin[5] on the roots of trinomial equations.

Our first lemma is

LEMMA 3.1. $\tau_1 \gtreqless \sigma_1 - 1$ *implies* $\tau_2 \lesseqgtr \sigma_2 - 1$.

This result follows at once from the fundamental relations.

Our second lemma is

LEMMA 3.2. *If* $(a_1, \cdots, a_N)$ *is a set of a's for which* $\sum_1^N a_n = S_0$ *is fixed and if* $S_1^{1/p_1} S_2^{1/p_2}$ *has an extreme value, then*

$$\sigma_2 A m^{\tau_2} a_m^{\sigma_2 - 1} + \sigma_1 B m^{\tau_1} a_m^{\sigma_1 - 1} = \kappa, \tag{3.1}$$

where $1 \leqq m \leqq N$, κ *is a positive parameter, and*

$$A = p_1 S_1^{1/p_1} S_2^{1/p_2 - 1}, \qquad B = p_2 S_1^{1/p_1 - 1} S_2^{1/p_2}.$$

The argument here is the same as that in Chapter II. The equations (3.1) may be written

$$m^{\tau_1} a_m^{\sigma_1 - 1} + \frac{\sigma_2 A}{\sigma_1 B} m^{\tau_2} a_m^{\sigma_2 - 1} = \frac{\kappa}{\sigma_1 B} \equiv k. \tag{3.2}$$

[4] G. H. Hardy, J. E. Littlewood, G. Pólya, *Inequalities*, Cambridge, 1934, p. 257.

[5] H. Mellin, *Zur Theorie der Trinomischen Gleichungen*, Annales Academiae Scientiarium Fennicae, series A, vol. 7(1916).

Let us multiply each equation by its root and sum from 1 to N. We get

$$S_1 + \frac{\sigma_2 A}{\sigma_1 B} S_2 = kS_0,$$

and since $A/B = p_1 S_1/p_2 S_2$,

$$S_1 = \frac{kS_0}{1 + \sigma_2 p_1/(\sigma_1 p_2)}.$$

We set $\alpha = S_1/S_2$, and, after some simple algebra, get

$$S_0 S_1^{-1/p_1} S_2^{-1/p_2} \equiv \left(1 + \frac{\sigma_2 p_1}{\sigma_1 p_2}\right)^{1/\sigma} \alpha^{1/p_2} k^{-(1/p_1+1/p_2)} S_0^{1/p_3}.$$

Hence we obtain

LEMMA 3.3. *The ratio is equal to a constant independent of N multiplied by*

$$\alpha^{1/p_2} k^{-1+1/p_3} S_0^{1/p_3},$$

with $0 < \alpha < \infty, 0 < k < \infty$.

Let us return to (3.2) and divide through by m^{τ_1}. Also put

$$\frac{\sigma_2 A}{\sigma_1 B} = \bar{\alpha},$$

and we get

$$a_m^{\sigma_1 - 1} + \bar{\alpha} m^{\tau_2 - \tau_1} a_m^{\sigma_2 - 1} = km^{-\tau_1}.$$

Now, if we put

$$a_m = m^{-\tau_1/(\sigma_1 - 1)} k^{1/(\sigma_1 - 1)} y_m, \tag{3.4}$$

the equation goes over into

$$y_m^{\sigma_1 - 1} + \{\bar{\alpha} k^{-(\sigma_1 - \sigma_2)/(\sigma_1 - 1)} m^{\delta}\} y_m^{\sigma_2 - 1} = 1,$$

where

$$\delta = \frac{\tau_1(1 - \sigma_2) - \tau_2(1 - \sigma_1)}{\sigma_1 - 1}.$$

We shall put

$$R = \bar{\alpha} k^{-(\sigma_1 - \sigma_2)/(\sigma_1 - 1)}, \qquad x = Rm^{\delta}, \tag{3.5}$$

and so,

$$y_m^{\sigma_1 - 1} + x y_m^{\sigma_2 - 1} = 1, \qquad 1 \leqq m \leqq N. \tag{3.6}$$

Hence we have

LEMMA 3.4. *The a's of Lemma 3.2 are given by (3.4) and the y's of this expression are the positive roots of (3.6), x being given by (3.5).*

We observe that these roots are all between zero and one since $x > 0$.

Next, we have the result:

LEMMA 3.5. *$\delta = 0$ implies, and is implied by, $\tau_1 = \sigma_1 - 1$. If $\tau_1 = \sigma_1 - 1$, there is no inequality.*

The first part of this lemma follows from the fundamental relations. The second part follows by an argument such as is used in Lemma 1.2 with $a_n = n^{-1}$ $(n = 1, 2, 3, \cdots)$.

We next prove

LEMMA 3.6. *Let y_m have the same meaning as in Lemma 3.4. Then for $1 \leqq m \leqq N$ we have*

$$y_m = \frac{1}{2\pi i} \cdot \frac{1}{\sigma_1 - 1} \int_{c-i\infty}^{c+i\infty} \frac{\Gamma(u)\Gamma[(1 - \overline{\sigma_2 - 1}u)(\sigma_1 - 1)^{-1}]}{\Gamma[(\sigma_1 + \overline{\sigma_1 - \sigma_2 u})(\sigma_1 - 1)^{-1}]} (Rm^{\delta})^{-u}\, du,$$

and $0 < c < (\sigma_2 - 1)^{-1}$.

In order to prove this we make use of a result of Mellin.[6] He shows that the branch of the function defined by

$$y^n + xy^{n_1} = 1, \qquad n > n_1,$$

which takes on the value 1 for $x = 0$ is given by

$$y(x) = \frac{1}{2\pi i} \cdot \frac{1}{n} \int_{c-i\infty}^{c+i\infty} \frac{\Gamma(u)\Gamma[(1 - n_1 u)n^{-1}]}{\Gamma[(1 + n + \overline{n - n_1 u})n^{-1}]} x^{-u}\, du,$$

where $0 < c < 1/n_1$. This result is valid in a certain angle containing the positive half of the real axis. Furthermore, the branch is real for $0 < x < \infty$ and $0 < y(x) < 1$.

Thus, we get Lemma 3.6 at once if we take

$$n = \sigma_1 - 1, \qquad n_1 = \sigma_2 - 1, \qquad x = Rm^{\delta}.$$

We now make use of Lemma 3.6 to compute

$$S_0 = \sum_1^N a_n = k^{1/(\sigma_1 - 1)} \sum_1^N y_m m^{-\tau_1/(\sigma_1 - 1)}$$

$$= \frac{k^{1/(\sigma_1-1)}}{2\pi i(\sigma_1 - 1)} \int_{c-i\infty}^{c+i\infty} \frac{\Gamma(u)\Gamma[(1 - \overline{\sigma_2 - 1}\,u)(\sigma_1 - 1)^{-1}]}{\Gamma[(\sigma_1 + \overline{\sigma_1 - \sigma_2 u})(\sigma_1 - 1)^{-1}]} R^{-u} \sum_1^N m^{-\delta u - \tau_1/(\sigma_1 - 1)}\, du,$$

where $0 < c < (\sigma_2 - 1)^{-1}$.

Now, we observe that the terms in the sum in the previous expression are all positive for every $0 < R < \infty$, since by the Mellin result they are positive roots of trinomial equations. Hence, we find that for $M \geqq N$,

$$S_0 \leqq \frac{k^{1/(\sigma_1-1)}}{2\pi i(\sigma_1 - 1)} \int_{c-i\infty}^{c+i\infty} b(u) R^{-u} \zeta_M\left(\delta u + \frac{\tau_1}{\sigma_1 - 1}\right) du,$$

[6] Op. cit.

where $0 < c < (\sigma_2 - 1)^{-1}$, where $b(u)$ is the ratio of gamma functions, and where

$$\zeta_M\left(\delta u + \frac{\tau_1}{\sigma_1 - 1}\right) = \sum_1^M m^{-\delta u - \tau_1/(\sigma_1 - 1)}.$$

Now, it is known that

$$\zeta_M\left(\delta u + \frac{\tau_1}{\sigma_1 - 1}\right) \to \zeta\left(\delta u + \frac{\tau_1}{\sigma_1 - 1}\right)$$

uniformly in any half-plane

$$\Re\left(\delta u + \frac{\tau_1}{\sigma_1 - 1}\right) \geqq 1 + \epsilon, \qquad \epsilon > 0.$$

We shall make use of this fact in the convergence proof we propose to give.

It will be necessary to break up our convergence proof into two parts, i.e., $\delta > 0$ and $\delta < 0$. We have already seen that when $\delta = 0$ there is no inequality of type H_1.

Let us consider the case $\delta < 0$ first. One easily sees that this implies

$$\tau_1 > \sigma_1 - 1, \qquad \tau_2 < \sigma_2 - 1,$$

and conversely. Henceforth, we shall refer to the case $\delta < 0$ as $\tau_1 > \sigma_1 - 1$. Next, the half-plane

$$\Re\left(\delta u + \frac{\tau_1}{\sigma_1 - 1}\right) > 1$$

is the half-plane

$$\Re(u) < \frac{1}{\delta}\left(1 - \frac{\tau_1}{\sigma_1 - 1}\right), \qquad \delta < 0.$$

We observe that the ratio of gamma functions has no poles in the strip

$$0 < \Re(u) < \frac{1}{\sigma_2 - 1}.$$

Finally, we show that

$$\frac{1}{\delta}\left\{1 - \frac{\tau_1}{\sigma_1 - 1}\right\} < \frac{1}{\sigma_2 - 1}.$$

Hence in the infinite strip

$$0 < \Re(u) < \frac{1}{\delta}\left\{1 - \frac{\tau_1}{\sigma_1 - 1}\right\} \tag{3.8}$$

our integrand is regular, and the series

$$\zeta_M\left(\delta u + \frac{\tau_1}{\sigma_1 - 1}\right)$$

converges uniformly in any strip in the interior of (3.8).

Let us now put

$$f(R) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} b(u)R^{-u}\zeta\left(\delta u + \frac{\tau_1}{\sigma_1 - 1}\right)du, \tag{3.9}$$

and suppose

$$\delta < 0, \qquad 0 < c < \frac{1}{\delta}\left\{1 - \frac{\tau_1}{\sigma_1 - 1}\right\}.$$

Under these conditions we shall prove

LEMMA 3.7. *If* $0 < R < \infty$, *then*

$$\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} b(u)R^{-u}\zeta_M\left(\delta u + \frac{\tau_1}{\sigma_1 - 1}\right)du \uparrow f(R), \qquad M \to +\infty.$$

We begin with the observation that this expression on the right in (3.9) actually defines a function $f(R)$ for $0 < R < \infty$. This will follow from the Mellin Inversion Theorem,[7] for[8]

$$b(u) = O(e^{-\pi|t|(\sigma_2-1)/(\sigma_1-1)}|t|^{-\frac{1}{2}}), \qquad |t| \to +\infty,$$

and

$$\zeta\left(\delta u + \frac{\tau_1}{\sigma_1 - 1}\right) = O(1)$$

on any line in the strip. Hence the function

$$F(u) = b(u)\zeta\left(\delta u + \frac{\tau_1}{\sigma_1 - 1}\right)$$

will satisfy the conditions of the inversion theorem, so that $f(x)$ will be analytic and regular in an angle containing the positive half of the real axis in the interior. In particular, $f(R)$ will be finite in any interval $\epsilon \leqq R \leqq \epsilon^{-1}$, $\epsilon > 0$. Furthermore, after we have established the convergence, it will follow that

$$f(R) > 0 \qquad (0 < R < \infty) \tag{3.10}$$

since the sequence in Lemma 3.7 is monotone increasing. The convergence follows easily from the uniform convergence of the ζ series in a half-plane. Hence we have

LEMMA 3.8.

$$S_0 \leqq \frac{k^{1/(\sigma_1-1)}}{\sigma_1 - 1}f(R).$$

We now make use of this result and Lemma 3.3 to get

$$\alpha^{1/p_2}k^{-1+1/p_3}S_0^{1/p_3} \leqq \alpha^{1/p_2}k^{-1+\sigma_1/[(\sigma_1-1)p_3]}\left\{\frac{f(R)}{\sigma_1 - 1}\right\}^{1/p_3}.$$

[7] H. Mellin's paper, p. 13.
[8] H. Mellin's paper. p. 15.

We set

$$-1 + \frac{1}{p_3}\frac{\sigma_1}{\sigma_1 - 1} \equiv A,$$

and may show that

$$A + \frac{\sigma_1 - \sigma_2}{\sigma_1 - 1}\frac{1}{p_2} \equiv 0.$$

Furthermore, α and $\bar{\alpha}$ differ by a constant multiplier independent of N. Hence we obtain

LEMMA 3.9.

$$S_0 S_1^{-1/p_1} S_2^{-1/p_2} \leqq \mathfrak{A} R^{1/p_2}\{f(R)\}^{1/p_3},$$

where $\mathfrak{A}$ *is independent of* N.

Thus, we have

LEMMA 3.10.

$$S_0 S_1^{-1/p_1} S_2^{-1/p_2} \leqq \mathfrak{A} \operatorname*{l.u.b.}_{0<R<\infty} [R^{1/p_2}\{f(R)\}^{1/p_3}],$$

valid for every N *and so for infinite series.*

We have already seen that

$$R^{1/p_2}\{f(R)\}^{1/p_3}$$

is continuous in any $\epsilon \leqq R \leqq \epsilon^{-1}$, $\epsilon > 0$. Hence we have reduced our problem to the proof of the following:

LEMMA 3.11. *If* $\tau_1 > \sigma_1 - 1$, *then* H_1 *exists if*

$$R^{1/p_2}\{f(R)\}^{1/p_3} = O(1),$$

as $R \to 0+$ *and as* $R \to +\infty$.

We shall prove

LEMMA 3.12.

$$f(R) = O(R^{-W}), \qquad R \to +\infty,$$

if

$$\tau_1 > \sigma_1 - 1, \qquad W = \frac{1}{\delta}\left\{1 - \frac{\tau_1}{\sigma_1 - 1}\right\}.$$

After having established this, we can make use of

$$\frac{1}{p_2} - W\frac{1}{p_3} \equiv 0,$$

and so obtain

LEMMA 3.13. *If* $\tau_1 > \sigma_1 - 1$, *then*

$$R^{1/p_2}\{f(R)\}^{1/p_3} = O(1), \qquad R \to +\infty.$$

Finally, the lemma that goes with Lemma 3.12 is the following:

LEMMA 3.14. *If* $\tau_1 > \sigma_1 - 1$, $f(R) = O(1)$, $R \to 0+$.

Hence we obtain the result

LEMMA 3.15. *If* $\tau_1 > \sigma_1 - 1$,

$$R^{1/p_2}\{f(R)\}^{1/p_3} = o(1), \qquad R \to 0+.$$

Thus, we shall have established the existence of H_1 for the case $\tau_1 > \sigma_1 - 1$ when we prove Lemmas 3.12 and 3.14.

We shall now prove Lemma 3.12. We begin by considering

$$\frac{1}{2\pi i}\oint F(z)x^{-z}\,dz \qquad (0 < x < \infty)$$

over the rectangle $ABCD$. Here BC is a segment of $\Re(z) = c$ and DA is a segment of $\Re(z) = h > c$. AB is above and CD is below the axis of reals. The poles of the integrand in this rectangle are those of

$$\Gamma\left(\frac{1 - (\sigma_2 - 1)z}{\sigma_1 - 1}\right), \tag{3.11}$$

and the one of

$$\zeta\left(\delta u + \frac{\tau_1}{\sigma_1 - 1}\right). \tag{3.12}$$

The pole of (3.12) is at the point

$$z = \frac{1}{\delta}\left\{1 - \frac{\tau_1}{\sigma_1 - 1}\right\} \equiv W,$$

and the residue is

$$\frac{1}{\delta}\,\Gamma\left(\frac{p_3}{p_2}\right)\Gamma\left(\frac{p_3}{p_1}\right)\{\Gamma(p_3)\}^{-1}x^{-W} \neq 0. \tag{3.13}$$

The poles of (3.11) in our rectangle are at the points

$$z_n = \frac{1 + n(\sigma_1 - 1)}{\sigma_2 - 1} \qquad (n = 0, 1, 2, \cdots, Q),$$

where Q is the largest integer in h. The residue at z_n is easily found to be

$$-\frac{\sigma_1 - 1}{\sigma_2 - 1}(-1)^n\,\frac{\Gamma(z_n)\zeta\left(\delta z_n + \dfrac{\tau_1}{\sigma_1 - 1}\right)x^{-z_n}}{\Gamma(1 + n)\Gamma\left(\dfrac{\sigma_1 + (\sigma_1 - \sigma_2)z_n}{\sigma_1 - 1}\right)}. \tag{3.14}$$

Hence, if we denote the coefficient of x in (3.13) by γ and the coefficient of x in (3.14) by γ_n we have

$$(3.15)\quad \frac{1}{2\pi i}\oint F(z)x^{-z}\,dz = \gamma\left(\frac{1}{x}\right)^{W} + \left(\frac{1}{x}\right)^{1/(\sigma_2-1)}\sum_{1}^{Q}\gamma_n\left(\frac{1}{x}\right)^{n(\sigma_1-1)/(\sigma_2-1)},$$

and the first term is the dominant one for large x.

Let us now consider the integrals along the lines AB and CD. As the lines tend to infinity, the first upwards and the second downwards, these integrals tend to zero. This follows from the order relation satisfied by the ratio of gamma functions and the order relation[9] for the function $\zeta(s)$.

Hence

$$\int_{h-i\infty}^{h+i\infty}(\quad)\,dz - \int_{c-i\infty}^{c+i\infty}(\quad)\,dz = \oint F(z)x^{-z}\,dz,$$

and so

$$(3.16)\qquad f(x) = -\frac{1}{2\pi i}\oint F(z)x^{-z}\,dz + \int_{h-i\infty}^{h+i\infty}(\quad)\,dz.$$

We shall now consider

$$(3.17)\qquad \left|\int_{h-i\infty}^{h+i\infty}F(z)x^{-z}\,dz\right| \leqq \int_{h-i\infty}^{h+i\infty}|F(z)|\,|x^{-z}|\,|dz|.$$

Recall that

$$|x^{-z}| = x^{-u} \qquad (0 < x < \infty)$$

and $u = \Re(z)$. Next $h = z_Q + \eta$, where

$$0 < \eta < \frac{\sigma_1 - 1}{\sigma_2 - 1},$$

and so (3.17) is equal to

$$(3.18)\qquad x^{-(z_Q+\eta)}\int_{h-i\infty}^{h+i\infty}|F(z)|\,|dz| = O(x^{-(z_Q+\eta)}),\qquad x\to+\infty.$$

Hence, from (3.16), (3.15) and (3.18) we have

$$f(x) = O(x^{-W}),\qquad x\to+\infty,$$

and this completes the proof of Lemma 3.12.

Finally, one can give a proof of Lemma 3.14 in the same manner. In this case we take a rectangle to the left of $\Re(z) = c$ instead of to the right as in the proof just given. This completes the existence proof for the case $\tau_1 > \sigma_1 - 1$.

[9] E. T. Whittaker and G. N. Watson, *Modern Analysis*, 3d edition, Cambridge, 1920, p. 270.

In the case $\tau_1 < \sigma_1 - 1$, we can argue in the same manner and prove

LEMMA 3.16. *The inequality of class H_1 will exist with a finite constant when $\tau_1 < \sigma_1 - 1$ if $R^{1/p_2}\{f(R)\}^{1/p_3} = O(1)$ as $R \to 0+$ and $R \to +\infty$.*

In this case, we define

$$f(R) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} b(u)R^{-u}\zeta\left(\delta u + \frac{\tau_1}{\sigma_1 - 1}\right) du,$$

and

$$\frac{1}{\delta}\left\{1 - \frac{\tau_1}{\sigma_1 - 1}\right\} < c < \frac{1}{\sigma_2 - 1}.$$

We prove by contour integration:

LEMMA 3.17. *If $\tau_1 < \sigma_1 - 1$,*

$$f(R) = O(R^{-W}), \qquad R \to 0+,$$

and

$$W = \frac{1}{\delta}\left\{1 - \frac{\tau_1}{\sigma_1 - 1}\right\}.$$

LEMMA 3.18. *If $\tau_1 < \sigma_1 - 1$,*

$$f(R) = O(R^{-1/(\sigma_2 - 1)}), \qquad R \to +\infty.$$

Next, for the function $R^{1/p_2}\{f(R)\}^{1/p_3}$, we get

LEMMA 3.19. *If $\tau_1 < \sigma_1 - 1$, then*

$$R^{1/p_2}\{f(R)\}^{1/p_3} = O(1), \qquad R \to 0+.$$

This follows from the fact that $1/p_2 - W/p_3 \equiv 0$.
For large $R > 0$ we get

LEMMA 3.20. *If $\tau_1 < \sigma_1 - 1$,*

$$R^{1/p_2}\{f(R)\}^{1/p_3} = o(1), \qquad R \to +\infty.$$

By Lemma 3.18, the exponent is

$$\frac{1}{p_2} - \frac{1}{p_3}\frac{1}{\sigma_2 - 1},$$

and this can be shown to be less than zero for every choice of the parameters.

The proof of Theorem 3.1 is now complete.

WASHINGTON, PENNSYLVANIA.

ANNIHILATORS OF QUADRATIC FORMS WITH APPLICATIONS TO PFAFFIAN SYSTEMS

By William G. McGavock

Introduction. This paper develops an algebraic approach to the study of certain arithmetic invariants of Pfaffian systems, thereby furnishing an extension of results previously obtained in connection with these invariants.[1] The principal algebraic result (Theorem 3.1) states that two quadratic forms defining a pencil of half-rank ρ in a Grassmann ring are simultaneously annihilated by the product of ρ linear forms. This result is employed to construct Pfaffian systems with half-rank ρ and species σ for all positive integers ρ, σ satisfying $\rho \leqq \sigma \leqq 2\rho$. This disproves a conjecture of Dearborn.[2] Finally we give a new upper bound for the species σ of a Pfaffian system of r equations, namely, $\sigma \leqq 2\rho + r - 1$.

1. **Pencils of forms.** By adjoining non-commutative marks $u_1, u_2, \cdots, u_n$ to a commutative field $\mathfrak{R}$ we obtain a Grassmann ring[3] which will be denoted by $\mathfrak{G}$.

Let S be a set of non-zero forms in $\mathfrak{G}$. S will be called a *pencil* if $a\omega + b\phi$ belongs to S whenever all the following three conditions are satisfied:

(i) a, b belong to $\mathfrak{R}$;

(ii) ω, ϕ belong to S;

(iii) $a\omega + b\phi \neq 0$.

The following properties of a pencil S follow directly from the definition of a pencil or are easily proved:

(a) Every member of S has non-negative degree.

(b) All members of S have the same degree.

(c) If S is a pencil, there is a positive integer r such that all members of S are given by

$$a_1\omega_1 + a_2\omega_2 + \cdots + a_r\omega_r,$$

where the a's range over $\mathfrak{R}$ independently, but are not simultaneously zero.

Received December 20, 1939; part of a doctoral dissertation presented June, 1939 at Duke University.

[1] See, for example, J. M. Thomas, *Pfaffian systems of species one*, Trans. Amer. Math. Soc., vol. 35(1933), pp. 356–371; Mabel Griffin, *Invariants of Pfaffian systems*, Trans. Amer. Math. Soc., vol. 35(1933), pp. 929–939; Donald Dearborn, *Inequalities among the invariants of Pfaffian systems*, this Journal, vol. 2(1936), pp. 705–711; J. M. Thomas, *A lower limit for the species of a Pfaffian system*, Proc. Nat. Acad. Sci., vol. 19(1933), p. 913.

[2] Loc. cit., p. 711.

[3] For a discussion of Grassmann algebra see J. M. Thomas, *Differential Systems*, New York, 1937, p. 10.

(d) Each form of a pencil has a unique representation for given ω's and minimum r.

The degree of S is defined to be the degree of any one of its forms.

The forms $\omega_1, \omega_2, \cdots, \omega_r$ will be called a *basis* for the pencil S, and r will be called the *dimension* of the pencil. The basis members are linearly independent.

Every finite or infinite set of forms of the same degree determines a pencil to which the forms belong.

2. **Annihilators and order.** If $F \neq 0$, $G \neq 0$, but $FG = 0$, then F is an *annihilator*[4] of G. The property of being an annihilator is reciprocal but not transitive. As a result of this definition we see that every Grassmann form of positive degree has annihilators.

A form G is an annihilator of a set $\mathfrak{F}$ of forms F_α $(\alpha = 1, 2, \cdots, r)$ if it is an annihilator of every form in the set.

If $GF = 0$, $HF = 0$, then $(aG + bH)F = 0$, even if a and b are simply interpreted as belonging to $\mathfrak{G}$ rather than as being restricted to $\mathfrak{R}$. Consequently, the totality of annihilators of F plus zero is an ideal in the Grassmann ring $\mathfrak{G}$. This ideal will be represented by $\mathfrak{G}_F$. $\mathfrak{G}_F$ always contains monomials since $u_1 u_2 \cdots u_n$ annihilates F_α.

The minimum degree of a monomial in $\mathfrak{G}_F$ is called the *order* of F. Clearly, the order $\leqq n$.

Interpreting a and b as members of $\mathfrak{R}$, we see that all annihilators of a given degree for a set of forms constitute a pencil.

3. **Reduction of two quadratic forms to forms with a common basis.** If S is a quadratic pencil of forms F_α, there is a positive integer ρ characterized by $F_{\alpha_1} F_{\alpha_2} \cdots F_{\alpha_{\rho+1}} = 0$ for every set of $\rho + 1$ α's and $F_{\alpha_1} F_{\alpha_2} \cdots F_{\alpha_\rho} \neq 0$ for at least one set $\alpha_1 \alpha_2 \cdots \alpha_\rho$. The name half-rank of the pencil will be applied to ρ.

If S has dimension one, it is well known[5] that ρ is the order of S. Cartan[6] has exhibited an example of a pencil which has dimension three, half-rank three and order four. Consequently, the order is not necessarily equal to the half-rank if the dimension exceeds one. This can also be established by the following simpler example,

$$F_1 = u_2 u_3, \qquad F_2 = u_3 u_1, \qquad F_3 = u_1 u_2$$

which has $\rho = 1$ and order 2.

The main theorem of this paper, now to be proved, will establish as a by-product that the order equals the half-rank if the dimension < 3.

[4] Name suggested by paper of Charles Hopkins, *Nil-rings with minimal condition for admissible left ideals*, this Journal, vol. 4(1938), pp. 664–667.

[5] See p. 31 of reference in footnote 3.

[6] See E. Cartan, *Sur les transformations de Bäcklund*, Bulletin de la Société Mathématique de France, vol. 43(1915), pp. 6–24; p. 13.

THEOREM 3.1. *If F and G are quadratic forms, if $\Omega = \omega_1 \cdots \omega_r$, where ω_i are linear, if $\Omega F^\lambda G^{\rho-\lambda+1} = 0$ ($\lambda = 0, 1, \cdots, \rho + 1$), and if $\Omega F^\rho \neq 0$, $\Omega G^\rho \neq 0$, then ΩF and ΩG have a common annihilator $u_1 u_2 \cdots u_\rho$; that is, we may write*

$$F = a_1\omega_1 + \cdots + a_r\omega_r + u_1v_1 + \cdots + u_\rho v_\rho ,$$

$$G = b_1\omega_1 + \cdots + b_r\omega_r + u_1w_1 + \cdots + u_\rho w_\rho .$$

Proof. We assume F, G can be written

$$(3.1) \quad F = a_1\omega_1 + \cdots + a_r\omega_r + u_1u_2 + \cdots + u_ju_{j+1} + \cdots + u_{2\lambda-1}u_{2\lambda} + f,$$

$$(3.2) \quad G = b_1\omega_1 + \cdots + b_r\omega_r + u_1u_{i_2} + \cdots + u_ju_l + \cdots + u_{2\lambda-1}u_{i_\lambda} + g,$$

$i_2, i_3, \cdots, i_\lambda$ being some arrangement of $2, 4, \cdots, 2\lambda - 2$, and hence

$$u_iu_1u_3 \cdots u_{2\lambda-1}f^{\rho-\lambda} \neq 0, \qquad u_iu_1u_3 \cdots u_{2\lambda-1}g^{\rho-\lambda} \neq 0,$$

for i distinct from $1, 3, \cdots, 2\lambda - 1$. We wish to show first that if f and g are not zero, then F, G can be written in forms *similar to* (3.1), (3.2), *but with* $r + \lambda + 1$ *terms explicitly displayed, or that F, G can be written directly* as forms with a common annihilator of degree ρ. If we can then show that F, G can actually be written in forms such as (3.1), (3.2), the theorem will be true.

Let $U = u_1u_3 \cdots u_{2\lambda-1}$. There are four possible cases:
(i) $u_{2\lambda}\Omega Ug^{\rho-\lambda} = u_l\Omega Uf^{\rho-\lambda} = 0$,
(ii) $u_{2\lambda}\Omega Ug^{\rho-\lambda} \neq 0$, $u_l\Omega Uf^{\rho-\lambda} \neq 0$,
(iii) $u_{2\lambda}\Omega Ug^{\rho-\lambda} = 0$, $u_l\Omega Uf^{\rho-\lambda} \neq 0$,
(iv) $u_{2\lambda}\Omega Ug^{\rho-\lambda} \neq 0$, $u_l\Omega Uf^{\rho-\lambda} = 0$.
Since (iii) and (iv) are entirely similar, we need examine only (i), (ii), (iii).

If (i) is satisfied, we have immediately

$$F = a_1\omega_1 + \cdots + a_r\omega_r + u_1u_2 + \cdots + u_ju_{j+1} + \cdots + u_{2\lambda-1}u_{2\lambda} + u_lu_{l+1} + \bar{f},$$

$$G = b_1\omega_1 + \cdots + b_r\omega_r + u_1u_{i_2} + \cdots + u_ju_l + \cdots + u_{2\lambda-1}u_{i_\lambda} + u_{2\lambda}u_{l+2} + \bar{g}.$$

If we note that u_ju_l, $u_{2\lambda}u_{l+2}$ may be paired with u_lu_{l+1}, $u_{2\lambda-1}u_{2\lambda}$ respectively and recall that $i_2, i_3, \cdots, i_\lambda$ is some arrangement of $2, 4, \cdots, 2\lambda - 2$, we see that F, G are in the required form, $l + 1$, and $l + 2$ playing the rôles of 2λ, l in (3.1), (3.2), respectively.

Suppose next that (ii) is satisfied. Assume that $\Omega f^{\rho-\alpha-\lambda+1}g^\alpha = 0$ for $\alpha \leqq p$. Consider $\Omega F^{\rho-p}G^{p+1}$. In this product the sum of those terms in which no member of U is absent will vanish separately from the remainder of the terms. This sum is of the form

$$\Omega(A_1f^{\rho-p-\lambda}g^{p+1} + A_2f^{\rho-p-\lambda+1}g^p + \cdots + A_{\lambda+1}f^{\rho-p}g^{p-\lambda+1}),$$

where A_j is a function of the u's. Since every term but the first vanishes by assumption, it too must vanish. Since $A_1 = u_1u_2 \cdots u_{2\lambda} \neq 0$ and since (ii) is true, it follows that $\Omega f^{\rho-p-\lambda}g^{p+1} = 0$. But by use of $\Omega F^{\rho+1} = 0$, it is readily seen that $\Omega f^{\rho-\lambda+1} = 0$. Hence by induction

$$\Omega f^{\rho-\alpha-\lambda+1}g^\alpha = 0 \qquad (\alpha = 0, 1, \cdots, \rho - \lambda + 1).$$

If the theorem is assumed true for forms H, K satisfying $\Omega H^{\alpha}K^{\beta} = 0$, where $\alpha + \beta < \rho + 1$, we may immediately write F, G as forms with a common basis, thus showing the theorem true for $\alpha + \beta = \rho + 1$.

But for $\rho = 1$ the theorem is immediate. Hence, the induction in this case is complete.

Finally we assume (iii) is true. Hence

$$F = a_1\omega_1 + \cdots + a_r\omega_r + u_1u_2 + \cdots + u_ju_{j+1} + \cdots + u_{2\lambda-1}u_{2\lambda} + f,$$

$$G = b_1\omega_1 + \cdots + b_r\omega_r + u_1u_{i_2} + \cdots + u_ju_l + \cdots + u_{2\lambda-1}u_{i_\lambda} + u_{2\lambda}\phi + \bar{g}.$$

If we have

$$\phi\Omega U f^{\rho-\lambda} = 0, \tag{3.3}$$

we may replace f by $\phi\psi + \bar{f}$ and we have the desired result. Hence, we assume $\phi\Omega U f^{\rho-\lambda} \neq 0$. Consider $\Omega F^{p+1}G^{\rho-p}$. Assume $\Omega f^{\alpha}\bar{g}^{\rho-\alpha-\lambda} = 0$ for $\alpha \leqq p$. In this product the sum of those terms from which no u displayed in G is absent must vanish independently of the other terms. This sum is of the form

$$\Omega A[f^{p-\lambda+1}\bar{g}^{\rho-p-1} + f^{p-\lambda+2}\bar{g}^{\rho-p-2} + \cdots + f^{p+1}\bar{g}^{\rho-p-\lambda-1}],$$

where $A = u_1u_{i_2} \cdots u_{2\lambda}\phi$. All these terms but the last vanish by assumption, and hence it also is zero. Since $A \neq 0$, it follows that $f^{p+1}\bar{g}^{\rho-p-\lambda-1} = 0$. But by use of $\Omega G^{\rho+1} = 0$, it is easily seen that $\Omega\bar{g}^{\rho-\lambda} = 0$. Hence, $\Omega f^{\alpha}\bar{g}^{\rho-\alpha-\lambda} = 0$ for $\alpha = 0, 1, \cdots, \rho - \lambda$. From this we infer $\Omega f^{\rho-\lambda} = 0$. This implies the contradiction $\Omega F^{\rho} = 0$. (iii) must therefore reduce to (3.3), (i) or (ii).

It remains to establish that F, G can be written in forms similar to (3.1), (3.2). Since $\Omega F^{\rho+1} = \Omega G^{\rho+1} = 0$, we may write

$$F = a_1\omega_1 + \cdots + a_r\omega_r + u_1u_2 + u_3u_4 + \cdots + u_{2\rho-1}u_{2\rho},$$

$$G = b_1\omega_1 + \cdots + b_r\omega_r + v_1v_2 + v_3v_4 + \cdots + v_{2\rho-1}v_{2\rho}.$$

$\Omega F^{\rho}G = 0$ gives $\Omega u_1u_2 \cdots u_{2\rho}v_1v_2 \cdots v_{2\rho} = 0$. This gives a linear relation of the type

$$\sum_{i=1}^{r} A_i\omega_i + \sum_{j=1}^{2\rho} B_ju_j = \sum_{j=1}^{2\rho} C_jv_j.$$

By virtue of this relation we may write

$$\begin{aligned} F &= a_1\omega_1 + \cdots + a_r\omega_r + u_1u_2 + f, \\ G &= b_1\omega_1 + \cdots + b_r\omega_r + u_1v_2 + g. \end{aligned} \tag{3.4}$$

Extending this one more step by the tactics employed above and adjusting notation we have

$$\begin{aligned} F &= a_1\omega_1 + \cdots + a_r\omega_r + u_1u_2 + u_3u_4 + \bar{f}, \\ G &= b_1\omega_1 + \cdots + b_r\omega_r + u_1u_3 + u_2u_5 + \bar{g}. \end{aligned} \tag{3.5}$$

Both (3.4) and (3.5) are of the form (3.1), (3.2). This completes the proof of the theorem.

An immediate consequence of this theorem is the

COROLLARY. *If the order exceeds the half-rank for a set of quadratic forms, then the dimension of the set exceeds two.*

4. **Reduction to canonical form.** In this section we give a method[7] of reducing two forms defining a quadratic pencil of dimension two simultaneously to canonical form. This process is analogous to a known reduction for a single form, although the passage from one to two forms presents some decided difficulties.

A quadratic form F of rank 2ρ can be written in canonical form[8]

$$F = U_1U_2 + \cdots + U_{2\rho-1}U_{2\rho}. \tag{4.1}$$

The U's appearing in this form will be called *canonical marks.* All the canonical marks are factors of F^ρ, and any linear factor of F^ρ can be made to play the rôle of U_1 in a reduction to canonical form. Moreover, any factor of F^ρ can be made one of the canonical marks for $F - U_1U_2$ provided it is not linearly dependent on U_1, U_2.

Suppose that two quadratic forms F, G each of which has rank $2r + 2\rho$ can be written simultaneously in the canonical forms

$$F = \omega + F_1, \qquad G = \varphi + G_1, \tag{4.2}$$

where

$$\begin{aligned} \omega &= a_1\omega_1 + \cdots + a_r\omega_r, \\ \varphi &= b_1\omega_1 + \cdots + b_r\omega_r, \end{aligned} \tag{4.3}$$

so that $F - F_1$, $G - G_1$ have the canonical marks $\omega_1, \cdots, \omega_r$ for a common basis. It follows by direct multiplication that

$$\begin{aligned} \omega_1 \cdots \omega_r F^\rho &\neq 0, & \omega_1 \cdots \omega_r G^\rho &\neq 0, \\ \omega_1 \cdots \omega_r F^{\rho+1} &= 0, & \omega_1 \cdots \omega_r G^{\rho+1} &= 0, \\ \omega_1 \cdots \omega_r F^\rho G &= 0, & \omega_1 \cdots \omega_r F G^\rho &= 0. \end{aligned} \tag{4.4}$$

Conversely, suppose (4.4) are satisfied by quadratic forms F, G and linear forms $\omega_1, \cdots, \omega_r$. Then[9] F, G can simultaneously be written in the form (4.2), where F_1, G_1 are in canonical form and ω, φ have the form (4.3). The two inequations in (4.4) show that (4.2) is a canonical form for F, G.

[7] This method of reduction, which incidentally furnishes an alternative proof of Theorem 3.1, was included at the suggestion of the referee.

[8] Cf. Theorem 17.1, p. 30 of the reference in footnote 3.

[9] See footnote 8.

Substitute from (4.2) in (4.4) and remove the factor $\omega_1 \cdots \omega_r$. We then have

$$F_1^\rho \neq 0, \qquad G_1^\rho \neq 0,$$
$$(4.5) \qquad F_1^{\rho+1} = 0, \qquad G_1^{\rho+1} = 0,$$
$$F_1^\rho G_1 = 0, \qquad F_1 G_1^\rho = 0.$$

The equation $F_1^\rho G_1 = 0$ shows that F_1, G_1 admit a canonical mark in common. By linear transformation of the marks this can be made u_1. Using u_1 as the first mark in reducing both F_1 and G_1 to canonical form we may write

$$(4.6) \qquad F_1 = u_1 v_1 + F_2, \qquad G_1 = u_1 w_1 + G_2,$$

where the canonical marks for F_2 form with u_1, v_1 an independent set and those for G_2 form with u_1, w_1 an independent set. Since F_2, G_2 are of rank $2\rho - 2$, we have

$$(4.7) \qquad \begin{array}{ll} F_2^{\rho-1} \neq 0, & G_2^{\rho-1} \neq 0, \\ F_2^\rho = 0, & G_2^\rho = 0. \end{array}$$

Use of these in (4.5) and a legitimate division by u_1 give

$$(4.8) \qquad v_1 F_2^{\rho-1} G_2 = 0, \qquad w_1 F_2 G_2^{\rho-1} = 0.$$

We wish to show that F_2, G_2 have a canonical mark in common. If either $F_2^{\rho-1}G_2 = 0$ or $F_2 G_2^{\rho-1} = 0$, then this follows at once, just as the existence of u_1 followed from the last line of (4.5). Consequently, we suppose

$$F_2^{\rho-1} G_2 \neq 0, \qquad F_2 G_2^{\rho-1} \neq 0.$$

The first of equations (4.8) then shows that some canonical mark w_2 for G_2 is a linear combination of v_1 and canonical marks for F_2. Hence we may write

$$w_2 = v_1 + u_2,$$

where u_2 is canonical for F_2 and w_2 for G_2. If F_2, G_2 are written in canonical form using u_2, w_2, we have

$$F_1 = u_1 v_1 + u_2 v_2 + \cdots, \qquad G_1 = u_1 w_1 + z_2(v_1 + u_2) + \cdots,$$

or rewriting

$$F_1 = u_1(v_1 + u_2) + u_2(v_2 + u_1) + \cdots,$$
$$G_1 = u_1 w_1 + z_2(v_1 + u_2) + \cdots.$$

Replacing $v_1 + u_2$ by $\bar{v}_1$, $v_2 + u_1$ by $\bar{v}_2$ and dropping the bars give

$$(4.9) \qquad F_1 = u_1 v_1 + u_2 v_2 + F_3, \qquad G_1 = u_1 w_1 + z_2 v_1 + G_3.$$

The second equation (4.8) can be applied to the modified F_2, G_2. It gives

$$w_1 F_2 G_2^{\rho-1} = 0.$$

Hence F_2 has a canonical mark, which can be made u_2, expressible in the form

$$u_2 = w_1 + z_3,$$

where z_3 is canonical for G_2. There are two possibilities: either

$$z_3 = az_2 + bv_1, \tag{4.10}$$

or z_3 is canonical for G_3.

If (4.10) holds, it is impossible for a to be zero because v_1 forms with the canonical marks of F_2 an independent set. Hence we may rewrite (4.9) as

$$F_1 = u_1(v_1 + au_1) + (w_1 + az_2 + bv_1)v_2 + \cdots,$$

$$G_1 = u_1(w_1 + az_2 + bv_1) + \left(z_2 + \frac{b}{a}v_1\right)(v_1 + au_1) + \cdots.$$

Set

$$\bar{v}_1 = v_1 + au_1,$$

$$\bar{w}_1 = w_1 + az_2 + bv_1,$$

$$\bar{z}_2 = z_2 + \frac{b}{a}v_1,$$

drop the bars and get

$$F_1 = u_1v_1 + w_1v_2 + F_3, \qquad G_1 = u_1w_1 + z_2v_1 + G_3. \tag{4.11}$$

Now consider the second possibility. When z_3 is canonical for G_3, write (4.9) as

$$F_1 = u_1v_1 + (w_1 + z_3)v_2 + \cdots,$$

$$G_1 = u_1(w_1 + z_3) + z_2v_1 + z_3(w_3 + u_1) + \cdots,$$

and thus reach the same result (4.11).

System (4.7) is readily seen to imply

$$\begin{aligned} &\omega_1 \cdots \omega_r v_1 F^{\rho-1} \neq 0, && \omega_1 \cdots \omega_r v_1 G^{\rho-1} \neq 0, \\ &\omega_1 \cdots \omega_r v_1 F^{\rho} = 0, && \omega_1 \cdots \omega_r v_1 G^{\rho} = 0, \\ &\omega_1 \cdots \omega_r v_1 F^{\rho-1} G = 0, && \omega_1 \cdots \omega_r v_1 F G^{\rho-1} = 0. \end{aligned} \tag{4.12}$$

These equations are the same as (4.4) except that the set $\omega_1, \cdots, \omega_r$ has been augmented by the linear form v_1 and the integer ρ has been decreased by unity. If the process is applied ρ times therefore, we have finally, with a slight change in notation, v_1 now being denoted by u_1,

$$\omega_1 \cdots \omega_r u_1 \cdots u_\rho F = 0, \qquad \omega_1 \cdots \omega_r u_1 \cdots u_\rho G = 0,$$

whence

$$F = \omega + F^*, \qquad G = \varphi + G^*,$$

where F^*, G^* are in canonical form with common annihilator $u_1u_2 \cdots u_\rho$.

THEOREM 4.1. *Two quadratic forms F, G satisfying* (4.4) *can be reduced simultaneously to canonical form. They have a common annihilator of $r + \rho$ linear forms among which are included $\omega_1, \cdots, \omega_r$.*

Consider now a pencil S of degree 2, of dimension 2 and of half-rank ρ. It is easy to see that the pencil has a basis F, G which satisfies

$$F^\rho \neq 0, \qquad G^\rho \neq 0, \qquad F^\alpha G^\beta = 0 \tag{4.13}$$

for all values of α, β satisfying $\alpha + \beta = \rho + 1$. In particular,

$$\begin{aligned} F^\rho &\neq 0, & G^\rho &\neq 0, \\ F^{\rho+1} &= 0, & G^{\rho+1} &= 0, \\ F^\rho G &= 0, & FG^\rho &= 0. \end{aligned} \tag{4.14}$$

Hence, letting the set of ω's be vacuous in (4.4), we see by Theorem 4.1 that F, G have a common annihilator of ρ linear forms.

5. Sets of inequalities for the algebraic case.

In this section we prove

THEOREM 5.1. *The order of a pencil of quadratic forms F_i does not exceed the rank of the pencil.*

Proof. Let 2ρ be the rank of the pencil. Then we may write

$$F_1 = u_1v_1 + u_2v_2 + \cdots + u_\rho v_\rho .$$

If F_j is one of the other forms of the pencil, we see that the u's and v's constitute a basis for it by virtue of the relation $F_1^\rho F_j = 0$. Since this is true of every other form of the pencil, the theorem follows.

We next construct examples in which the order assumes all values from ρ to 2ρ inclusive.

Consider the set

$$\begin{aligned} F &= u_1u_2 + u_4u_5 + u_7u_8 + \cdots + u_{3\rho-2}u_{3\rho-1}, \\ G &= u_2u_3 + u_5u_6 + u_8u_9 + \cdots + u_{3\rho-1}u_{3\rho}, \\ H &= u_3u_1 + u_6u_4 + u_9u_7 + \cdots + u_{3\rho}u_{3\rho-2}, \end{aligned} \tag{5.1}$$

where

$$u_1u_2u_3 \cdots u_{3\rho-2}u_{3\rho-1}u_{3\rho} \neq 0. \tag{5.2}$$

We first prove that $F^iG^jH^k = 0$ if $i + j + k = \rho + 1$. Write

$$F = f + u_{3n-2}u_{3n-1}, \qquad G = g + u_{3n-1}u_{3n}, \qquad H = h + u_{3n}u_{3n-2}.$$

Assume $f^\alpha g^\beta h^\gamma = 0$ for $\alpha + \beta + \gamma = n$.

$$\begin{aligned} F^iG^jH^k &= [f^i + \lambda f^{i-1}u_{3n-2}u_{3n-1}][g^j + \mu g^{j-1}u_{3n-1}u_{3n}][h^k + \nu h^{k-1}u_{3n}u_{3n-2}] \\ &= f^ig^jh^k + \nu f^ig^jh^{k-1}u_{3n}u_{3n-2} + \mu f^ig^{j-1}h^ku_{3n-1}u_{3n} + \lambda f^{i-1}g^jh^ku_{3n-2}u_{3n-1}. \end{aligned}$$

But since $i + j + k = n + 1$, it results that these four terms vanish by virtue of our assumption. It is readily verified that for $n = 2$ we have $F^\alpha G^\beta H^\gamma = 0$, where $\alpha + \beta + \gamma = 3$. Hence, the induction is complete.

The modification of (5.1) that we shall use in the subsequent treatment will leave this result invariant (see (5.3)). For example, by assuming the truth of the statement for

$$
\begin{aligned}
F &= f + u_1u_2 + \cdots + u_{3i-2}u_{3i-1}, \\
G &= g + u_1u_3 + \cdots + u_{3i-2}u_{3i}, \\
H &= h + u_1u_{\lambda_1} + \cdots + u_{3i-2}u_{\lambda_i},
\end{aligned}
$$

where f, g, h are the parts of F, G, H following the scheme of (5.1), we can readily show that it holds when we replace i by $i + 1$. It is then easy to verify that the statement holds for

$$F = f + u_1u_2, \qquad G = g + u_1u_3, \qquad H = h + u_1u_{\lambda_1}.$$

In the annihilator of (5.1) there must be one or more factors of the type $a_1u_1 + a_2u_2 + \cdots + a_{3\lambda}u_{3\lambda}$ in which a_1, a_2, a_3 are not all zero. Clearly one such factor is insufficient, for $Au_1u_2 = Au_2u_3 = Au_3u_1 = 0$, where $A = a_1u_1 + a_2u_2 + \cdots + a_{3\lambda}u_{3\lambda}$, implies $a_1 = a_2 = a_3 = 0$. Two factors on the other hand are sufficient, as for example u_1, u_2. Similar reasoning is employed for the group of second terms, third terms, etc. In view of (5.2) we are therefore led to the conclusion that the order $= 2\rho$.

Consider the following modification of (5.1).

$$
\begin{aligned}
&F = u_1u_2 + u_4u_5 + \cdots + u_{3i-2}u_{3i-1} + u_{3i+1}u_{3i+2} + \cdots + u_{3\rho-2}u_{3\rho-1}, \\
(5.3)\quad &G = u_1u_3 + u_4u_6 + \cdots + u_{3i-2}u_{3i} + u_{3i+2}u_{3i+3} + \cdots + u_{3\rho-1}u_{3\rho}, \\
&H = u_1u_{\lambda_1} + u_4u_{\lambda_2} + \cdots + u_{3i-2}u_{\lambda_i} + u_{3i+3}u_{3i+1} + \cdots + u_{3\rho}u_{3\rho-2}, \\
&\qquad u_1u_2 \cdots u_{3\rho}u_{\lambda_1}u_{\lambda_2} \cdots u_{\lambda_i} \neq 0 \ (i = 0, 1, \cdots, \rho).
\end{aligned}
$$

A linear annihilator reduces all first terms to zero. Similarly for all second terms, all i-th terms. As in (5.1) we see that a quadratic annihilator is necessary and sufficient to use for each of the last $\rho - i$ groups of terms. Hence, order $= 2\rho - 2i + i = 2\rho - i$. We thus see that the order may have any value between ρ and 2ρ inclusive.

The sets of inequalities

$$(5.4) \qquad r \leqq 2, \qquad \text{order} = \rho;$$

$$(5.5) \qquad 3 \leqq r \leqq k, \qquad \rho \leqq \text{order} \leqq 2\rho$$

are such that every non-negative integral solution of either comprises invariants of some pencil of quadratic forms. k may clearly be taken to be at least 3ρ.

6. **Pfaffian systems.** In the subsequent discussion we interpret $\mathfrak{R}$ as a field of functions of n independent variables $x_1, \cdots, x_n$ closed under differentiation.

We take $u = dx$. A member of $\mathfrak{G}$ is a differential form or a member of $\mathfrak{R}$. If a member of $\mathfrak{G}$ is a linear differential form, it is called a *Pfaffian*. A pencil of Pfaffians is called a *Pfaffian system*. If the dimension of this pencil is r, we may choose any r independent members of the pencil as a basis for the Pfaffian system.

Let a_i belong to $\mathfrak{R}$. Representing a Pfaffian by ω, we have $\omega = a_i\,dx_i$. Associated with every linear differential form ω of $\mathfrak{G}$ there is a quadratic differential form of $\mathfrak{G}$ called the derivative of ω. If we represent this form by ω', we have $\omega' = da_i\,dx_i$.

7. **Pfaffian systems with species satisfying** $\rho \leqq \sigma \leqq 2\rho$. By means of the algebraic theory developed above we can now construct a Pfaffian system with $r > 2$ basis members, half-rank ρ, species satisfying $\rho \leqq \sigma \leqq 2\rho$, and with class $p = 5\rho + r - \sigma$.

Consider the following pencil of dimension three which will be shown to have half-rank ρ and whose species can be made to be ρ, $\rho + 1$, $\cdots$, or 2ρ.

$$
\begin{aligned}
\omega_1 &= x_1\,dx_2 + x_4\,dx_5 + \cdots + x_{3i-2}\,dx_{3i-1} + x_{3i+1}\,dx_{3i+2} \\
&\qquad\qquad + \cdots + x_{3\rho-2}\,dx_{3\rho-1} + dx_{\lambda_{i+1}}, \\
\omega_2 &= x_1\,dx_3 + x_4\,dx_6 + \cdots + x_{3i-2}\,dx_{3i} + x_{3i+2}\,dx_{3i+3} \\
&\qquad\qquad + \cdots + x_{3\rho-1}\,dx_{3\rho} + dx_{\lambda_{i+2}}, \\
\omega_3 &= x_1\,dx_{\lambda_1} + x_4\,dx_{\lambda_2} + \cdots + x_{3i-2}\,dx_{\lambda_i} + x_{3i+3}\,dx_{3i+1} \\
&\qquad\qquad + \cdots + x_{3\rho}\,dx_{3\rho-2} + dx_{\lambda_{i+3}}, \\
&dx_1\,dx_2 \cdots dx_{3\rho}\,dx_{\lambda_1} \cdots dx_{\lambda_i} \cdots dx_{\lambda_{i+3}} \neq 0.
\end{aligned}
\tag{7.1}
$$

That (7.1) has half-rank ρ is evident from the facts that $\Omega(\omega_1)'^{\rho}$ contains the non-zero term $dx_1\,dx_2 \cdots dx_{3\rho-2}\,dx_{3\rho-1}\,dx_{3\rho}\cdot dx_{\lambda_{i+1}}\,dx_{\lambda_{i+2}}\,dx_{\lambda_{i+3}}$ which can be canceled by no other term in the product, and that all products in the ω_α' of degree $\rho + 1$ must vanish since in (5.3) all products of degree $\rho + 1$ vanish.

Comparing the derived system of (7.1) with (5.3), we see that the order $=$ $2\rho - i$ for it. But any form of degree $2\rho - i$ whose product by ω_α' is zero will necessarily be a form in the differentials dx_1, dx_2, $\cdots$, $dx_{3\rho}$ alone. It is therefore impossible for ω_1, ω_2, or ω_3 to be a factor of such a form due to the presence of $dx_{\lambda_{i+1}}$, $dx_{\lambda_{i+2}}$, $dx_{\lambda_{i+3}}$, respectively. Hence, $2\rho - i$ forms must be adjoined to (7.1). It is clear that these may be chosen as differentials. For example, we might adjoin dx_1, dx_4, $\cdots$, dx_{3i-2}, dx_{3i+1}, dx_{3i+2}, $\cdots$, $dx_{3\rho-2}$, $dx_{3\rho-1}$. We have shown that (7.1) is of species $2\rho - i$ ($i = 0, 1, \cdots$, or ρ).

It follows immediately that given any $r > 2$, any ρ, and any σ satisfying $\rho \leqq \sigma \leqq 2\rho$, we can construct a Pfaffian system satisfying these conditions. For example, adjoin dx_j ($j = \lambda_{i+4}, \cdots, \lambda_{i+r}$) to (7.1), where

$$dx_1\,dx_2 \cdots dx_{3\rho} \cdots dx_{\lambda_1}\,dx_{\lambda_2} \cdots dx_{\lambda_i}\,dx_{\lambda_{i+1}} \cdots dx_{\lambda_{i+r}} \neq 0.$$

Since (7.1) satisfies $r = 3$, $\rho = \rho$, $\sigma = 2\rho - i$ $(i = 0, 1, \cdots, \rho)$, it follows that the new system satisfies

$$r = r, \qquad \rho = \rho, \qquad \sigma = 2\rho - i$$

since the addition of dx_j affects neither σ nor ρ.

We now prove that for the augmented system (7.1) the class is $p = 5\rho + r - \sigma$.

In this case the associated set of the derived forms $(\omega_1)'$, $(\omega_2)'$, $(\omega_3)'$ will be included in the characteristic system. From this it is easy to deduce that the characteristic system has as basis

$$dx_1, dx_2, \cdots, dx_{3\rho}, dx_{\lambda_{i+1}}, \cdots, dx_{\lambda_{i+r}}.$$

Hence the class $p = 3\rho + r + i$. But $\sigma = 2\rho - i$. Consequently, $p = 5\rho + r - \sigma$.

We thus have Pfaffian systems corresponding to each solution in non-negative integers of the inequalities

$$r = r, \qquad \rho \leqq \sigma \leqq 2\rho, \qquad p = 5\rho + r - \sigma.$$

8. An upper bound for the species.

Dearborn[10] has obtained $\rho(r + 1) - 1$ as an upper bound for the species of a Pfaffian system. Furthermore, he has completely discussed the inequalities of systems having $\rho = 1$. It is our purpose now to determine for the species an upper bound less than $\rho(r + 1) - 1$ for all values of $\rho > 1$, $r > 1$. We shall prove

THEOREM 8.1. *The species of a Pfaffian system with r basis members does not exceed* $2\rho + r - 1$.

Proof. In Theorem 3.1 let $(\omega_1)' = F$ and $(\omega_2)' = G$, $\omega_1 = \omega_1, \omega_2 = \omega_2, \cdots, \omega_r = \omega_r$. We may then write

$$(\omega_1)' = a_{11}\omega_1 + a_{12}\omega_2 + \cdots + a_{1r}\omega_r + u_1u_2 + u_3u_4 + \cdots + u_{2\rho-1}u_{2\rho},$$

$$(\omega_2)' = a_{21}\omega_1 + a_{22}\omega_2 + \cdots + a_{2r}\omega_r + u_1\alpha_1 + u_3\alpha_3 + \cdots + u_{2\rho-1}\alpha_{2\rho-1},$$

$$(\omega_3)' = a_{31}\omega_1 + a_{32}\omega_2 + \cdots + a_{3r}\omega_r + \beta_{31}\beta_{32} + \beta_{33}\beta_{34} + \cdots + \beta_{3,2\rho-1}\beta_{3,2\rho},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$(\omega_r)' = a_{r1}\omega_1 + a_{r2}\omega_2 + \cdots + a_{rr}\omega_r + \beta_{r1}\beta_{r2} + \beta_{r3}\beta_{r4} + \cdots + \beta_{r,2\rho-1}\beta_{r,2\rho}.$$

By use of Theorem 5.1 we may write

$$(\omega_i)' = a_{i1}\omega_1 + a_{i2}\omega_2 + \cdots + a_{ir}\omega_r + \sum_{k=1}^{\rho}\gamma_{ik}\sum_{j=1}^{2\rho}C_{i,jk}u_j \qquad (i = 3, 4, \cdots, r),$$

the C's being functions of the α's. By factoring out u_1, u_2, etc. and making a linear transformation on their coefficients, we may write

$$(\omega_i)' = a_{i1}\omega_1 + a_{i2}\omega_2 + \cdots + a_{ir}\omega_r + w_{i1}u_1 + w_{i2}u_2 + \cdots + w_{i,2\rho}u_{2\rho}.$$

[10] See footnote 2.

Now consider the two Pfaffian systems with basis members

$$\omega_1, \omega_2, \omega_3, \omega_4, \cdots, \omega_r,$$

and

$$\omega_1, \cdots, \omega_r, u_1, \cdots, u_{2\rho}, a_{12}, \cdots, a_{1r}. \tag{8.1}$$

We shall show that the system with (8.1) as basis is passive. For this purpose form $(\omega_1')'$,

$$(\omega_1')' = (a_{11})'\omega_1 - a_{11}(\omega_1)' + \cdots + (a_{1r})'\omega_r - a_{1r}(\omega_r)' + (u_1)'u_2 - u_1(u_2)' + \cdots + (u_{2\rho-1})'u_{2\rho} - u_{2\rho-1}(u_{2\rho})' = 0.$$

Put $\Omega = \omega_1 \cdots \omega_r u_1 \cdots u_{2\rho} a_{12} \cdots a_{1r}$. Clearly $\Omega(\omega_j)' = 0$ $(j = 1, 2, \cdots, r)$. We prove now that $\Omega(u_i)' = 0$ $(i = 1, 2, \cdots, 2\rho)$. Suppose i is even. Multiply $(\omega_1')'$ by $\omega_1 \cdots \omega_r u_1 \cdots u_{i-2} u_i \cdots u_{2\rho} a_{12} \cdots a_{1r}$, getting $-\omega_1 \cdots \omega_r u_1 \cdots u_{i-2} u_i \cdots u_{2\rho} a_{12} \cdots a_{1r}(-a_{11}[\omega_1]' - u_{i-1}[u_i]') = 0$. The first term is zero and hence we have $\Omega(u_i)' = 0$. Obviously when i is odd the same treatment holds.

Finally we prove $\Omega(a_{1j})' = 0$ $(j = 1, 2, \cdots, r)$. Multiply $(\omega_1')'$ by $\omega_1 \cdots \omega_{j-1}\omega_{j+1} \cdots \omega_r u_1 \cdots u_{2\rho} a_{12} \cdots a_{1r}$, getting $\omega_1 \cdots \omega_{j-1}\omega_{j+1} \cdots \omega_r u_1 \cdots u_{2\rho} a_{12} \cdots a_{1i}(-a_{11}[\omega_1]' + [a_{1j}]'\omega_j) = 0$. Again the first term is zero and consequently we have $\Omega(a_{1j})' = 0$. Hence (8.1) is passive and the theorem is established.

Duke University.

CYCLIC TRANSITIVITY

By T. Radó and P. Reichelderfer

Introduction and fundamental concepts

0.1. Let us denote by 1 a set which will serve as our space; the elements of the set 1 will be called points. However, we shall not assume that the set 1 is topologized in any way; that is, 1 is a wholly unconditioned set, unless a statement to the contrary is explicitly made.

0.2. Given in 1 a binary relation $\mathfrak{R}$, we shall write $a \mathfrak{R} b$ to express the fact that the points a and b of 1 are in the $\mathfrak{R}$-relation. Many important binary relations arising in algebra are reflexive, symmetric, and transitive; that is, $a \mathfrak{R} a$ for every point a; $a \mathfrak{R} b$ implies $b \mathfrak{R} a$ for every pair of points a, b; and $a \mathfrak{R} b \mathfrak{R} c$ implies $a \mathfrak{R} c$ for every triple of points a, b, c. On the other hand, the general theory of sets leads to binary relations—such as set inclusion—which are transitive, but are neither reflexive nor symmetric.[1] Binary relations of the types just mentioned have been studied and applied extensively. Both of these types are transitive. In this paper we are concerned with binary relations which are reflexive and symmetric, but are not necessarily transitive; the requirement of transitivity is replaced by a weaker condition which we shall call *cyclic transitivity*, and which we now describe.

0.3. Given a binary relation $\mathfrak{R}$ in 1, we say that $\mathfrak{R}$ is *cyclically transitive* if, for every finite cyclically ordered set of distinct points $a_1, a_2, \cdots, a_n$ satisfying $a_1 \mathfrak{R} a_2 \mathfrak{R} \cdots \mathfrak{R} a_n \mathfrak{R} a_1$, we have $a_i \mathfrak{R} a_j$ for every choice of the subscripts i and j. Let $\mathfrak{R}$ be a reflexive and symmetric binary relation; if $\mathfrak{R}$ is transitive (cf. 0.2), then clearly $\mathfrak{R}$ is cyclically transitive, but the converse is not true. Thus cyclic transitivity is an extension of ordinary transitivity, that is, an extension of one of the fundamental concepts arising in algebra. On the other hand, we shall see presently (cf. 0.4) that cyclic transitivity also arises in connection with certain fundamental concepts in topology.

Received February 26, 1940; presented to the American Mathematical Society, December 26, 1939. This is a condensed version of our original paper which was accepted for publication by the Fundamenta Mathematicae in August, 1939. In this paper we tried to arrange the definitions, lemmas, and theorems in such an order that the reader may construct the proofs for himself with the aid of hints given. Explicit proofs are given only in a few cases where the proof depends upon a device which might not readily occur to the reader.

[1] For an extensive and detailed study of transitive relations, see, for example, Foradori [1], [2], [3]. (Numbers in square brackets indicate references in the bibliography at the end of this paper.) We want to thank Professor Rainich for these references.

0.4. Let us now assume that 1 is a set sufficiently topologized so that we can speak of mutually separated sets, hence of connected sets in 1. A basic concept in paper [4] by K. W.[2] is that of conjugate points. Two (not necessarily distinct) points a and b are conjugate if, for every choice of the point x different from a and b, the points a and b are in the same component of $1 - x$. Writing $a \mathfrak{R} b$ to mean that a and b are conjugate, we easily verify that $\mathfrak{R}$ is reflexive and symmetric, but not generally transitive. However, if we have $a \mathfrak{R} b_1 \mathfrak{R} c$, $a \mathfrak{R} b_2 \mathfrak{R} c$, where b_1 and b_2 are distinct points, then it follows easily that $a \mathfrak{R} c$.[3] A closer inspection of the properties of $\mathfrak{R}$ reveals that this latter property is but a special case of the more general property of cyclic transitivity (cf. 0.3) possessed by $\mathfrak{R}$. In paper [6] Whyburn uses the concept of a nodular set as a basis for his work. A nodular set S is a connected set which is disconnected by none of its points; that is, $S - x$ is connected for every choice of x. A nodular set is called a maximal nodular set if it is a proper subset of no nodular set; a non-degenerate maximal nodular set is called a nodule. Whyburn considers no binary relation explicitly; however, we may associate a binary relation with the concepts in Whyburn [6] as follows: $a \mathfrak{R} b$ if either $a = b$ or a and b are on the same nodule. Inspection reveals that $\mathfrak{R}$ is a reflexive, symmetric, and cyclically transitive binary relation which is not generally transitive. Moore, in paper [5], uses the notion of two points being separated by a third point as a basic concept. Two points a and b are said to be separated by a point x if $a \neq x \neq b$, and if $1 - x$ is a sum of two mutually separated sets A and B such that $a \in A$, $b \in B$. Consider the following binary relation: $a \mathfrak{R} b$ if and only if no point x different from a and b separates a and b in the sense of Moore. Obviously $\mathfrak{R}$ is reflexive and symmetric, but is not generally transitive. Again, a closer inspection shows that $\mathfrak{R}$ is cyclically transitive.

In so far as we are aware, the cyclic transitivity property of the binary relations so intimately related to these three theories in topology has not been stated or used explicitly. Yet, once attention is called to this property, it is quite apparent that cyclic transitivity accounts for many of the fundamental results in the theories just mentioned.

0.5. Thus the concept of cyclic transitivity may be construed to have its origin both in algebra and topology (cf. 0.2). The purpose of this paper is to study the concept of cyclic transitivity in a manner suggested by its dual origin. How does cyclic transitivity arise? We now delineate two important ways of generating all possible reflexive, symmetric, and cyclically transitive binary relations in 1; the second of these is developed and discussed more fully in Chapter I.

[2] The symbol K. W. is used consistently to refer jointly to C. Kuratowski and G. T. Whyburn.

[3] This property was studied in an unpublished paper by T. Radó, of which this paper is an improvement and an extension. For an abstract, see the Bulletin of the American Mathematical Society, vol. 45(1939), p. 373.

A first mode of generation is the following. Let there be given, for every point x in the space 1, a binary relation $\mathfrak{R}_x$ which is defined in $1 - x$ and is reflexive, symmetric, and transitive there. Define in 1 a binary relation as follows: $a\, \mathfrak{R}(\mathfrak{R}_x)\, b$ if and only if $a\, \mathfrak{R}_x\, b$ for every choice of x different from a and b. The relation $\mathfrak{R}(\mathfrak{R}_x)$ is clearly reflexive, symmetric, and cyclically transitive; conversely, as the reader will easily verify, every reflexive, symmetric, and cyclically transitive binary relation can be generated in this way.

A second mode of generation is obtained as follows. Let Γ be a class of subsets of 1 possessing the following properties:

Property $\mathfrak{P}_1$: The empty set 0, the whole space 1, and every set consisting of a single point of 1, is in Γ.

Property $\mathfrak{P}_2$: If Ω is any subclass of Γ such that the product of all the sets in Ω is not empty, then the sum of all the sets in Ω is a set in Γ.

Given now any set S, we define a Γ-component of S to be a maximal set with respect to the property of being both a subset of S and a set in Γ. From properties $\mathfrak{P}_1$ and $\mathfrak{P}_2$ it follows that S is the sum of its Γ-components, and two distinct Γ-components of S have no point in common.

Next, we define a binary relation as follows: $a\, \mathfrak{R}(\Gamma)\, b$ if and only if, for every choice of x different from a and b, the points a and b are in the same Γ-component of $1 - x$. This relation $\mathfrak{R}(\Gamma)$ is clearly reflexive, symmetric, and cyclically transitive; conversely, every reflexive, symmetric, and cyclically transitive binary relation can be generated in this way (cf. 1.7).

0.6. In fact, there are generally several classes possessing properties $\mathfrak{P}_1$ and $\mathfrak{P}_2$ and generating the same binary relation (cf. 1.10). A general example of this, important for the sequel, is the following one. Given any class Γ possessing properties $\mathfrak{P}_1$ and $\mathfrak{P}_2$, define a class Γ' as follows: a set S belongs to Γ' if it is either one of the sets described in property $\mathfrak{P}_1$ (cf. 0.5) or if, for every choice of the point x in $1 - S$, the set S is in one Γ-component of $1 - x$. We shall call Γ' the *closure* of Γ. The class Γ' possesses properties $\mathfrak{P}_1$ and $\mathfrak{P}_2$, and generates the same binary relation as does Γ; that is, $\mathfrak{R}(\Gamma') = \mathfrak{R}(\Gamma)$ (cf. 1.9). A class Γ possessing properties $\mathfrak{P}_1$ and $\mathfrak{P}_2$ is said to be *closed* if $\Gamma = \Gamma'$ (cf. 1.9). A necessary and sufficient condition that a class Γ be closed is that Γ possess the following additional

Property $\mathfrak{P}_3$: If Ω^* is any subclass of Γ, then the product of all the sets in Ω^* is a set in Γ.

0.7. We next consider a system consisting of a reflexive, symmetric, and cyclically transitive binary relation $\mathfrak{R}$, any one of the classes Γ generating $\mathfrak{R}$ (cf. 0.5), and its closure Γ' (which also generates $\mathfrak{R}$ (cf. 0.6))—briefly, a system ($\mathfrak{R}$, Γ, Γ'). Such a system gives rise to a sequence of concepts and theorems which correspond closely to those arising in the theories referred to in 0.4. For example, consider the cyclic element theory in Peano spaces (cf. K. W. [4]). One of the fundamental concepts of that theory is the concept of a cyclic chain

joining two distinct points a and b in the Peano space: it is the product of all $\mathfrak{A}$-sets containing the points a and b, where an $\mathfrak{A}$-set is defined to be a closed set which contains every arc whose end points are in the set. In our untopologized space 1 (cf. 0.1) this definition would be meaningless. It is quite interesting then to observe that we are able to define for the untopologized space 1 a concept which reduces to that of a cyclic chain when 1 is a Peano space, and for which theorems similar to those for cyclic chains in Peano spaces are valid (cf. 1.19, 1.20). To stress this analogy, we call the concept which we introduce a cyclic chain. Similar remarks apply to nearly all the fundamental concepts and the terminology used in this paper; for details, the reader may consult the sections in Chapter I from 1.13 onward.

0.8. As we specified in 0.1, our space 1 is wholly untopologized in the usual sense. However, the introduction of the class of sets Γ (cf. 0.5) does topologize 1 in a fashion—the sets in Γ correspond to connected sets in the standard treatments (cf. 0.4). This observation suggests the desirability of a full axiomatic treatment of the theory of the structure of a general space, if we use the notion of a "connected" set as an undefined concept. Such a treatment is beyond the scope of this paper.

0.9. It is also beyond the scope of this paper to discuss applications of the abstract theory herein developed. However, it is not difficult to verify that, in the important special case when 1 is a Peano space and the class Γ (cf. 0.5) is the class of all connected sets in 1, our concepts are equivalent to those in K. W. [4], although our definitions for these concepts necessarily differ considerably from theirs. Curiously enough, it seems uneconomical to try directly to identify our concepts with those in K. W. [4]; rather, it is easy to develop the structure theorems in that theory from our general theory, and then our concepts are automatically identified with theirs. This was explained in some detail in the original version of this paper.

Cyclically transitive binary relations

1.1. Until further notice, $\mathfrak{R}$ will denote a binary relation which is reflexive, symmetric, and cyclically transitive (cf. 0.3). Given $\mathfrak{R}$, we define various concepts which depend solely upon $\mathfrak{R}$. A set S is called *coherent* if it is non-degenerate and if every two points of S are in the $\mathfrak{R}$-relation; that is, $a \in S$, $b \in S$ imply $a \mathfrak{R} b$ (cf. 0.2). Clearly every non-degenerate subset of a coherent set is coherent. A set S is called *complete* if it is non-degenerate and if it contains every point which is in the $\mathfrak{R}$-relation to two distinct points of S; that is, $a \mathfrak{R} x \mathfrak{R} b$, $a \in S$, $b \in S$, $a \neq b$, imply $x \in S$. It is evident that the product of complete sets, if it is non-degenerate, is complete. A set is called a *proper cyclic element* if it is both coherent and complete; the letter C is used consistently in the sequel to denote a proper cyclic element. Clearly a proper cyclic element is a maximal coherent set, that is, a non-degenerate set which is maximal with respect to the

property that any two elements of the set are in the $\mathfrak{R}$-relation (cf. 1.23). Thus a proper cyclic element in our theory corresponds to a fundamental concept in algebra, namely, that of a residue class. If $\mathfrak{R}$ were also transitive, then, indeed, our proper cyclic elements would be residue classes in the set 1. The product of any two residue classes is empty; in our theory there is a corresponding result for proper cyclic elements which we presently state.

1.2. **Theorem.** *If two proper cyclic elements C_1 and C_2 have more than one point in common, they are identical.*

The proof follows from the definition of a proper cyclic element, using only the reflexivity and the symmetry of $\mathfrak{R}$.

1.3. **Theorem.** *Given two distinct points a and b satisfying $a \mathfrak{R} b$, there exists one and only one proper cyclic element containing them.*

The existence part of the proof is obtained by verifying that the (non-degenerate) set S containing all points $x \in 1$ such that $a \mathfrak{R} x \mathfrak{R} b$ is both coherent and complete. In contradistinction to the proof of the theorem in 1.2, this proof depends essentially upon the cyclic transitivity of $\mathfrak{R}$. The uniqueness part of the proof follows directly from 1.2.

1.4. A finite ordered set of distinct points $a_1, a_2, \cdots, a_n$ satisfying $a_1 \mathfrak{R} a_2 \mathfrak{R} \cdots \mathfrak{R} a_n$ is called an *$\mathfrak{R}$-chain* joining a_1 and a_n. Each of the points $a_1, a_2, \cdots, a_n$ is called a vertex of the $\mathfrak{R}$-chain; in particular, the points $a_2, \cdots, a_{n-1}$ are called interior vertices. Clearly an ordered pair of distinct points a, b is an $\mathfrak{R}$-chain joining a and b if and only if $a \mathfrak{R} b$. Similarly, a *closed $\mathfrak{R}$-chain* is a finite cyclically ordered set of distinct points $a_1, a_2, \cdots, a_n$ satisfying $a_1 \mathfrak{R} a_2 \mathfrak{R} \cdots \mathfrak{R} a_n \mathfrak{R} a_1$. Of course, in the case of a closed $\mathfrak{R}$-chain there is no occasion to speak of interior vertices, since each vertex plays the same rôle. The fact that $\mathfrak{R}$ is cyclically transitive may now be expressed in the following equivalent form: any two vertices of a closed $\mathfrak{R}$-chain are in the $\mathfrak{R}$-relation (cf. 0.3). Consider two distinct points a and b such that there is an $\mathfrak{R}$-chain joining them. Then, in the class of all $\mathfrak{R}$-chains joining a and b, there is obviously at least one *minimal* $\mathfrak{R}$-chain, that is, one with the smallest possible number of vertices. In fact, there exists exactly one minimal $\mathfrak{R}$-chain joining a and b (cf. 1.5).

Lemma. *If $\mathfrak{C}_0$ is a minimal $\mathfrak{R}$-chain joining two distinct points a and b, if $\mathfrak{C}$ is any $\mathfrak{R}$-chain joining a and b, then every vertex of $\mathfrak{C}_0$ is also a vertex of $\mathfrak{C}$.*

The proof, which is left for the reader, depends essentially upon the cyclic transitivity of $\mathfrak{R}$.

1.5. As an easy consequence of the lemma in 1.4, we have the

Theorem. *If a and b are two distinct points such that there exists an $\mathfrak{R}$-chain joining them, then there exists exactly one minimal $\mathfrak{R}$-chain joining a and b.*

1.6. With the given binary relation $\mathfrak{R}$ (cf. 1.1) we associate a class of sets $\Lambda(\mathfrak{R})$ defined as follows: 0, 1, every point in 1 is in $\Lambda(\mathfrak{R})$; a non-degenerate set E is in $\Lambda(\mathfrak{R})$ if, for every pair of distinct points a and b in E, there exists an $\mathfrak{R}$-chain joining a and b all of whose vertices are in E. Obviously $\Lambda(\mathfrak{R})$ possesses properties $\mathfrak{P}_1$ and $\mathfrak{P}_2$ (cf. 0.5). Moreover, we have the

LEMMA. *The set* $\Lambda(\mathfrak{R})$ *possesses property* $\mathfrak{P}_3$ (cf. 0.6).

The proof follows from the definition of $\Lambda(\mathfrak{R})$ by use of the lemma in 1.4.

1.7. THEOREM. *Given a binary relation* $\mathfrak{R}^*$, *there exists a class possessing properties* $\mathfrak{P}_1$ *and* $\mathfrak{P}_2$ *and generating* $\mathfrak{R}^*$ (cf. 0.5) *if and only if* $\mathfrak{R}^*$ *is reflexive, symmetric, and cyclically transitive.*

If $\mathfrak{R}^*$ is reflexive, symmetric, and cyclically transitive, then the reader will easily verify that the class $\Lambda(\mathfrak{R}^*)$ defined in 1.6 possesses properties $\mathfrak{P}_1$, $\mathfrak{P}_2$ (and $\mathfrak{P}_3$) and generates $\mathfrak{R}^*$. Next, if Γ be any class possessing properties $\mathfrak{P}_1$ and $\mathfrak{P}_2$ and generating $\mathfrak{R}^*$, then $\mathfrak{R}^*$ is obviously reflexive and symmetric; the cyclic transitivity of $\mathfrak{R}^*$ follows easily by the use of closed $\mathfrak{R}^*$-chains (cf. 1.4).

1.8. Given a reflexive, symmetric, and cyclically transitive binary relation $\mathfrak{R}$, how many classes do we have which possess properties $\mathfrak{P}_1$ and $\mathfrak{P}_2$ and generate $\mathfrak{R}$? We have exhibited one in 1.7, viz., $\Lambda(\mathfrak{R})$. We shall show that generally there are several.

1.9. LEMMA. *If* Γ *is any class possessing properties* $\mathfrak{P}_1$ *and* $\mathfrak{P}_2$ *and generating the binary relation* $\mathfrak{R}(\Gamma)$ (cf. 0.5), *then its closure* Γ' (cf. 0.6) *possesses properties* $\mathfrak{P}_1$, $\mathfrak{P}_2$, $\mathfrak{P}_3$, *and also generates* $\mathfrak{R}(\Gamma)$; *that is,* $\mathfrak{R}(\Gamma') = \mathfrak{R}(\Gamma)$.

The proof is obvious.

LEMMA. *A class* Γ *possessing properties* $\mathfrak{P}_1$ *and* $\mathfrak{P}_2$ *is closed if and only if it also possesses property* $\mathfrak{P}_3$ (cf. 0.6).

Proof. That $\mathfrak{P}_3$ is a property of Γ if Γ is closed is a consequence of the preceding lemma. Conversely, suppose that Γ is a class possessing properties $\mathfrak{P}_1$, $\mathfrak{P}_2$, and $\mathfrak{P}_3$. We assert that $\Gamma = \Gamma'$. For, if E is any set in Γ' (cf. 0.6), which is not one of the sets described in $\mathfrak{P}_1$, then for every point $x \in 1 - E$, the set E is in a Γ-component S_x of $1 - x$. Now clearly $E = \Pi S_x$ for $x \in 1 - E$. Thus $E \in \Gamma$, since Γ possesses property $\mathfrak{P}_3$. So $\Gamma' \subset \Gamma$, but we always have $\Gamma \subset \Gamma'$. Consequently $\Gamma = \Gamma'$; that is, Γ is closed.

1.10. Let Γ be any class possessing properties $\mathfrak{P}_1$ and $\mathfrak{P}_2$, but not property $\mathfrak{P}_3$. Then Γ and its closure Γ' are two distinct classes generating the same reflexive, symmetric, and cyclically transitive binary relation $\mathfrak{R} = \mathfrak{R}(\Gamma) = \mathfrak{R}(\Gamma')$ (cf. 1.9). Also $\Lambda(\mathfrak{R})$ (cf. 1.6) is a closed class generating the binary relation $\mathfrak{R}$ (cf. 1.8). Thus we see three generally distinct classes possessing properties $\mathfrak{P}_1$ and $\mathfrak{P}_2$ and generating the same binary relation $\mathfrak{R}$. The outstanding example

where these three classes are generally distinct is furnished by the study of cyclic elements in a Peano space (cf. 1.13). Note that Γ' depends solely upon the class Γ of which it is the closure (cf. 0.6); on the other hand, the class $\Lambda(\mathfrak{R})$ depends solely upon the binary relation $\mathfrak{R}$ (cf. 1.6). Moreover, the class $\Lambda(\mathfrak{R})$ is unique in the sense that it is the smallest closed class generating the binary relation $\mathfrak{R}$; this fact is a consequence of the

LEMMA. *If Γ is any closed class, that is, a class possessing properties $\mathfrak{P}_1$, $\mathfrak{P}_2$, and $\mathfrak{P}_3$, which generates the binary relation $\mathfrak{R} = \mathfrak{R}(\Gamma)$, then the class $\Lambda(\mathfrak{R})$ (cf. 1.6) is a subclass of Γ.*

Proof. Suppose $E \in \Lambda(\mathfrak{R})$. If E is 0, 1, or consists of a single point, then $E \in \Gamma$ by property $\mathfrak{P}_1$ (cf. 0.5). Otherwise, let a and x be two distinct points of E; denote by $\mathfrak{C}(x)$ an $\mathfrak{R}$-chain joining a and x all of whose vertices are in E (cf. 1.6). Suppose y is any point in $1 - \mathfrak{C}(x)$; then $\mathfrak{C}(x)$ is in one Γ-component S_y of $1 - y$ (cf. 0.5). Now $\mathfrak{C}(x) = \Pi S_y$ for $y \in 1 - \mathfrak{C}(x)$, so $\mathfrak{C}(x) \in \Gamma$ by property $\mathfrak{P}_3$ (cf. 0.6). But $E = \sum \mathfrak{C}(x)$ for $x \in E - a$; hence $E \in \Gamma$ by $\mathfrak{P}_2$ (cf. 0.5). Thus $\Lambda(\mathfrak{R}) \subset \Gamma$.

1.11. Since there is a smallest closed class generating a given reflexive, symmetric, and cyclically transitive binary relation $\mathfrak{R}$ (cf. 1.10), we naturally ask: Is there a smallest class possessing properties $\mathfrak{P}_1$ and $\mathfrak{P}_2$ and generating $\mathfrak{R}$? Is there a largest class possessing properties $\mathfrak{P}_1$ and $\mathfrak{P}_2$ and generating $\mathfrak{R}$? Simple examples, which will occur to the reader after a little thought, show that generally the answers to both of these questions are in the negative.

1.12. By way of illustration, we point out, for a Peano space 1, four important classes which are generally distinct, possess properties $\mathfrak{P}_1$ and $\mathfrak{P}_2$, and generate the same reflexive, symmetric, and cyclically transitive binary relation. Let Γ be the class comprised of all the connected sets in 1, including 0, 1, and every set consisting of a single point in 1. Let Γ' be the closure of Γ (cf. 0.6). Now Γ and Γ' possess properties $\mathfrak{P}_1$ and $\mathfrak{P}_2$ and generate the same binary relation $\mathfrak{R} = \mathfrak{R}(\Gamma) = \mathfrak{R}(\Gamma')$ (cf. 1.9). Also $\Lambda(\mathfrak{R})$ is a class possessing properties $\mathfrak{P}_1$ and $\mathfrak{P}_2$ (and $\mathfrak{P}_3$), and generating $\mathfrak{R}$ (cf. 1.10). But there is another important class which possesses properties $\mathfrak{P}_1$ and $\mathfrak{P}_2$, and which generates $\mathfrak{R}$—namely, the class comprised of 0, 1, every set consisting of a single point in 1, and every connected *open* set in 1.

1.13. Given a class of sets Γ possessing properties $\mathfrak{P}_1$ and $\mathfrak{P}_2$ (cf. 0.5); it gives rise to a closed class Γ' possessing properties $\mathfrak{P}_1$, $\mathfrak{P}_2$, and $\mathfrak{P}_3$ (cf. 0.6). We have seen that Γ and Γ' generate the same binary relation $\mathfrak{R} = \mathfrak{R}(\Gamma) = \mathfrak{R}(\Gamma')$, which is reflexive, symmetric, and cyclically transitive (cf. 1.9). In the rest of this chapter we regard Γ as fixed, and introduce for study various concepts depending upon one or more of the entities Γ, Γ', $\mathfrak{R}$ (cf. 1.14, 1.15, 1.17, 1.19).

1.14. A point x is called a *cut point* of 1 if $1 - x$ is a set not in Γ. A point x is called an *end point* if it is not a cut point and there exists no point $y \epsilon 1$ distinct from x such that $y \mathfrak{R} x$ (cf. 0.2). A point x is said to *separate* two distinct points a and b if $a \neq x \neq b$, and a and b are in different Γ-components of $1 - x$ (cf. 0.5). Denote by $K(a, b)$ the totality of points in 1 each of which separates a and b. The following statements are now obvious. Every point in $K(a, b)$ is a cut point. The set $K(a, b)$ is empty if and only if $a \mathfrak{R} b$ (cf. 0.5). If a and b are any two distinct points, and if $a + b \subset E$ where $E \epsilon \Gamma'$, then $K(a, b) \subset E - (a + b)$. It follows that $a + K(a, b) + b \epsilon \Gamma'$. Finally, every coherent set (cf. 1.1) is in Γ'; thus every proper cyclic element (cf. 1.1) is in Γ', although the simplest examples show that a proper cyclic element is not generally in Γ.

1.15. A configuration consisting of $n > 1$ distinct points $a_1, \cdots, a_n$, and n sets $S_1, \cdots, S_n$ satisfying the conditions: $S_i \epsilon \Gamma'$ for $i = 1, \cdots, n$; $S_1 \cdot S_2 = a_2$, $S_2 \cdot S_3 = a_3, \cdots, S_n \cdot S_1 = a_1$; $S_i \cdot S_j = 0$ for $1 < |i - j| < n - 1$ and $i, j = 1, \cdots, n$ is called a *closed polygon*. Such a configuration is denoted by the symbol $(a_1, \cdots, a_n; S_1, \cdots, S_n)$. A closed polygon is a generalization of a closed $\mathfrak{R}$-chain (cf. 1.4) in the following sense: if the finite cyclically ordered set of distinct points $a_1, \cdots, a_n$ satisfies $a_1 \mathfrak{R} a_2 \mathfrak{R} \cdots \mathfrak{R} a_n \mathfrak{R} a_1$—that is, constitutes a closed $\mathfrak{R}$-chain—then the configuration $(a_1, a_2, \cdots, a_n; a_1 + a_2, a_2 + a_3, \cdots, a_n + a_1)$ is clearly a closed polygon. The points $a_1, \cdots, a_n$ are called the vertices of the polygon; the sets $S_1, \cdots, S_n$ are called the sides of the polygon. The two vertices on a side of a polygon are called adjacent vertices.

1.16. Theorem. *If $(a_1, \cdots, a_n; S_1, \cdots, S_n)$ is a closed polygon, then the set $\sum_{i=1}^{n} a_i$ is coherent.*

The proof may be made by showing that any two adjacent vertices are in the $\mathfrak{R}$-relation since $\mathfrak{R}$ is cyclically transitive (cf. 1.13).

From this theorem we have the easily proved corollaries:

Corollary 1. *If the $n > 2$ distinct points $a_1, \cdots, a_n$, together with the set $S \epsilon \Gamma'$, satisfy the relations $a_i \mathfrak{R} a_{i+1}$ $(i = 1, \cdots, n - 1)$ and $S \cdot \sum_{i=1}^{n} a_i = a_n + a_1$, then the set $\sum_{i=1}^{n} a_i$ is coherent.*

Corollary 2. *If S_1 and S_2 are two sets satisfying $S_1 \epsilon \Gamma'$, $S_2 \epsilon \Gamma'$, $S_1 \neq 0 \neq S_2$, $S_1 \cdot S_2 = 0$, then there is at most one proper cyclic element C such that $S_1 \cdot C \neq 0 \neq S_2 \cdot C$.*

Definition. Let S be any non-degenerate set; denote by $\Psi(S)$ the class of all proper cyclic elements C each of which has at least two distinct points in common with S—that is, for which $S \cdot C$ is non-degenerate.

Corollary 3. *Given a non-degenerate set S and a non-empty set S^* satisfying $S \epsilon \Gamma'$, $S^* \epsilon \Gamma'$, $S \cdot S^* = 0$, there exists at most one proper cyclic element $C \epsilon \Psi(S)$ such that $C \cdot S^* \neq 0$.*

Corollary 4. *If x is not a cut point* (cf. 1.14), *then there exists at most one proper cyclic element containing x.*

Corollary 5. *Given two distinct proper cyclic elements C_1 and C_2 in $\Psi(S)$, where S is any non-degenerate set in Γ'; either C_1 and C_2 have no points in common, or else they have a single point of S in common, which is a cut point.*

1.17. We introduce further notions needed in the sequel. A set which is complete (cf. 1.1) and belongs to the class Γ' is called an $\mathfrak{H}$-set. The letter H is used consistently to denote an $\mathfrak{H}$-set. By definition, every $\mathfrak{H}$-set is non-degenerate. Clearly the product of $\mathfrak{H}$-sets, if it is non-degenerate, is an $\mathfrak{H}$-set (cf. 1.1, 1.13). If S is any non-degenerate set, the product of all $\mathfrak{H}$-sets containing S is evidently the smallest $\mathfrak{H}$-set containing S—denote it by $H(S)$. We now state a property of $\mathfrak{H}$-sets, which the reader may easily verify by use of closed polygons (cf. 1.15).

Lemma. *If H_1 and H_2 are two $\mathfrak{H}$-sets such that $H_1 \cdot H_2$ consists of a single point, then $H_1 + H_2$ is an $\mathfrak{H}$-set.*

Extending this result, we have the

Corollary. *If $H_1, \cdots, H_n$ are any finite number of $\mathfrak{H}$-sets such that every H_k for $k = 2, \cdots, n$ has exactly one point in common with the set $\sum_{j=1}^{k-1} H_j$, then $\sum_{j=1}^{n} H_j$ is an $\mathfrak{H}$-set.*

1.18. Regarding the structure of $\mathfrak{H}$-sets, we have the

Theorem. *If S is any non-degenerate set in Γ', then*

$$H(S) = S + \sum_{C \epsilon \Psi(S)} C.$$

The proof of this theorem will follow easily from two facts concerning the set

$$(*) \qquad E = S + \sum_{C \epsilon \Psi(S)} C.$$

Lemma 1. *The set E of* (*) *is an $\mathfrak{H}$-set.*

The proof follows without difficulty from Corollary 1 in 1.16.

Lemma 2. *Let S and E have meanings as above. If H is any $\mathfrak{H}$-set such that $H \cdot E$ is non-degenerate, then $H \cdot S$ is non-degenerate.*

The proof follows readily from the theorem and Corollary 3 in 1.16.

From Lemma 2 we have the

COROLLARY. *Let S and E have meanings as above. If C is any proper cyclic element such that $C \cdot E$ is non-degenerate, then $C \epsilon \Psi(S)$.*

The proof of the theorem stated at the beginning of this section is now immediate.

Replacing E of (*) by $H(S)$ in Lemma 2 and its corollary, we have the

THEOREM. *Let S be any non-degenerate set in Γ'. If H is any $\mathfrak{H}$-set such that $H \cdot H(S)$ is non-degenerate, then $H \cdot S$ is non-degenerate. If C is any proper cyclic element such that $C \cdot H(S)$ is non-degenerate, then $C \epsilon \Psi(S)$.*

1.19. If a and b are two distinct points, the $\mathfrak{H}$-set $H(a + b)$ is called the *cyclic chain* joining a and b (cf. 1.17). We denote this cyclic chain by $C(a, b)$. The following theorems are readily established.

THEOREM. *For any two distinct points a and b, we have*

$$C(a, b) = H(a + K(a, b) + b).$$

THEOREM. *Given two distinct points a and b, a necessary and sufficient condition that the cyclic chain $C(a, b)$ be a proper cyclic element is that $a \mathfrak{R} b$.*

1.20. Given two distinct points a and b. Since the set $a + K(a, b) + b$ is in Γ' (cf. 1.14), we can apply the results of 1.18 to discuss the structure of the cyclic chain $C(a, b)$ (cf. 1.19). The cyclic chain $C(a, b)$ is the sum of sets in Γ', viz., the sets $a + K(a, b) + b$, and every proper cyclic element having two or more distinct points in common with $a + K(a, b) + b$; if a proper cyclic element is not in $C(a, b)$, it has at most one point in common with $C(a, b)$. If C_1 and C_2 are two distinct proper cyclic elements in $C(a, b)$, they have at most one point in common; if they have a point in common, it is a point of the set $a + K(a, b) + b$, and is a cut point (cf. Corollary 5, 1.16). If $K(a, b)$ is empty, that is, if $a \mathfrak{R} b$ (cf. 1.14), then $C(a, b)$ is the unique proper cyclic element containing a and b. Finally, if H is any $\mathfrak{H}$-set having just one point in common with $a + K(a, b) + b$, then that point is the only point which H and $C(a, b)$ have in common (cf. 1.18).

1.21. When 1 is a Peano space, it is true that, if two distinct proper cyclic elements C_1 and C_2 in a cyclic chain $C(a, b)$ have a point in common, that point belongs to $K(a, b)$. However, simple examples show that generally we can only assert that this common point is in the set $a + K(a, b) + b$ (cf. 1.20); that is, it may be the point a or b.

1.22. From the definition (cf. 1.1) it is clear that a proper cyclic element depends solely upon the binary relation $\mathfrak{R}$ and not upon the particular class which generates $\mathfrak{R}$. Further, if S is any non-degenerate set, then the class $\Psi(S)$ of all proper cyclic elements each of which has two or more distinct points in common with S is quite independent of the choice of the class which generates $\mathfrak{R}$ (cf. 1.16). Now simple examples show that generally the cyclic chain depends

upon the particular class which generates $\mathfrak{R}$ (cf. 1.20). However, we do have the following

THEOREM. *If a and b are two distinct points, and if there is an $\mathfrak{R}$-chain* (cf. 1.4) *joining a and b, then the cyclic chain $C(a, b)$ is independent of the class which generates $\mathfrak{R}$.*

The proof is made by observing that the unique minimal $\mathfrak{R}$-chain joining a and b (cf. 1.5) is simply the set $a + K(a, b) + b$. Then it is clear from the structure of $C(a, b)$ that $C(a, b)$ is independent of the choice of the class which generates $\mathfrak{R}$ (cf. 1.20).

1.23. If we generalize the definition of a proper cyclic element found in K. W. [4], we have the

DEFINITION. A proper cyclic element is the set of all points each of which is in the $\mathfrak{R}$-relation to some point x which is neither a cut point nor an end point (cf. 1.14).

It is clear that every proper cyclic element according to the generalized K. W. definition is also a proper cyclic element according to our definition, for, as we remarked in 1.1, our proper cyclic element is a non-degenerate maximal coherent set. But the converse is not true for a general space 1, as we shall see presently. By definition, a proper cyclic element in the generalized K. W. sense must always contain at least one point of the space which is not a cut point, and hence which is in no other proper cyclic element (cf. Corollary 4, 1.16). We now give an example in which every point of a certain proper cyclic element in our sense is a cut point, since it belongs to another proper cyclic element.

Example. Let U denote the closed linear interval $v = 0$, $0 \leqq u \leqq 1$, and V_u the closed linear interval $u = u$, $0 \leqq v \leqq 1$. Our space 1 shall be the set $U + \sum V_u$ for $u \in U$. Consider the binary relation: $a \mathfrak{R} b$ if and only if a and b are on U, or on the same V_u. The reader will verify that $\mathfrak{R}$ is reflexive, symmetric, and cyclically transitive. Each of the sets U and V_u, $0 \leqq u \leqq 1$, is evidently a proper cyclic element, and there are no others. Consider the proper cyclic element U. Every point $(u_0, 0)$, $0 \leqq u_0 \leqq 1$, of U is a cut point, since it belongs to two distinct proper cyclic elements, viz., U and V_{u_0} (cf. Corollary 4, 1.16). And this statement is valid, no matter what class is regarded to be the generator of $\mathfrak{R}$ (cf. 0.5).

This example shows another feature of the general theory which differs from that of the K. W. theory. For a general space we may have a non-denumerable number of distinct proper cyclic elements. However, for the Peano space considered in K. W. [4], the number of distinct proper cyclic elements is always denumerable.

BIBLIOGRAPHY

1. E. FORADORI, *Grundbegriffe einer allgemeinen Teiltheorie*, Monatshefte für Mathematik und Physik, vol. 39(1932), pp. 439–454.

2. E. FORADORI, *Stetigkeit und Kontinuität als Teilbarkeitseigenschaften*, Monatshefte für Mathematik und Physik, vol. 40(1933), pp. 161–180.
3. E. FORADORI, *Das endliche Mass*, Monatshefte für Mathematik und Physik, vol. 41(1934), pp. 133–173.
4. C. KURATOWSKI AND G. T. WHYBURN, *Sur les éléments cycliques et leurs applications*, Fundamenta Mathematicae, vol. 16(1930), pp. 305–331.
5. R. L. MOORE, *Foundations of Point Set Theory*, American Mathematical Society Colloquium Publications, vol. 13, 1932, p. 62.
6. G. T. WHYBURN, *On the structure of connected and connected im kleinen point sets*, Transactions of the American Mathematical Society, vol. 32(1930), pp. 926–943.

OHIO STATE UNIVERSITY.

A SET OF POLYNOMIALS

BY L. CARLITZ

1. Let $GF(p^n)$ denote a Galois (finite) field of order p^n. Let M denote a polynomial in an indeterminate x with coefficients in $GF(p^n)$:

$$M = M(x) = c_0x^m + c_1x^{m-1} + \cdots + c_m\,;$$

for $c_0 \neq 0$, we write deg $M = m$; for $c_0 = 1$, M is called primary. Further let[1]

$$\psi_m(t) = \prod_{\deg M < m} (t - M), \qquad \psi_0(t) = t, \tag{1.1}$$

where t is another indeterminate and the product extends over all M (including 0) of degree $< m$; then we have the formula

$$\psi_m(t) = \sum_{i=0}^{m} (-1)^{m-i} \begin{bmatrix} m \\ i \end{bmatrix} t^{p^{ni}}, \tag{1.2}$$

where

$$\begin{bmatrix} m \\ i \end{bmatrix} = \frac{F_m}{F_i L_{m-i}^{p^{ni}}}, \qquad \begin{bmatrix} m \\ 0 \end{bmatrix} = \frac{F_m}{L_m}, \qquad \begin{bmatrix} m \\ m \end{bmatrix} = 1, \tag{1.3}$$

and

$$\begin{aligned} F_m &= [m][m-1]^{p^n} \cdots [1]^{p^{n(m-1)}}, & F_0 &= 1, \\ L_m &= [m][m-1] \cdots [1], & L_0 &= 1, \\ [m] &= x^{p^{nm}} - x. \end{aligned} \tag{1.4}$$

We remark that

$$\psi_m(x^m) = \psi_m(M) = F_m\,,$$

for M primary of degree m, so that F_m is the product of the primary polynomials of degree m.

For k an arbitrary integer $\geqq 0$, put

$$k = \alpha_0 + \alpha_1 p^n + \cdots + \alpha_s p^{ns} \qquad (0 \leqq \alpha_i < p^n), \tag{1.5}$$

and define the polynomial g_k by means of

$$g_k = F_1^{\alpha_1} \cdots F_s^{\alpha_s}, \qquad g_0 = 1. \tag{1.6}$$

Received April 10, 1940.

[1] See this Journal, *On certain functions connected with polynomials in a Galois field*, vol. 1(1935), pp. 137–168, p. 141. This paper will be cited as I.

Properties of g_k have been discussed in a previous paper.[2] It is evident that g_k may be thought of as a generalization of F_m. We shall now similarly generalize $\psi_m(t)$; we define

$$(1.7)\qquad G_k(t) = \psi_0^{\alpha_0}(t)\psi_1^{\alpha_1}(t)\cdots\psi_s^{\alpha_s}(t), \qquad G_0(t) = 1.$$

Thus in particular

$$(1.8)\qquad G_{\alpha p^{ni}}(t) = \psi_i^{\alpha}(t) \qquad (0 \leqq \alpha < p^n).$$

Note that $G_k(t)$ is of degree k in t, and that the coefficients of the powers of t are polynomials in x. In this paper we shall derive various properties of $G_k(t)$.

It will be convenient to define another polynomial, also of degree k in t, which is closely related to $G_k(t)$; we put

$$(1.9)\qquad G'_k(t) = \prod_{i=0}^{s} G'_{\alpha_i p^{ni}}(t),$$

where α_i is defined by (1.5), and

$$(1.10)\qquad G'_{\alpha p^{ni}}(t) = \begin{cases} \psi_i^{\alpha} & \text{for } 0 \leqq \alpha < p^n - 1, \\ \psi_i^{\alpha} - F_i^{\alpha} & \text{for } \alpha = p^n - 1. \end{cases}$$

In particular for

$$k = p^{nm} - 1 = (p^n - 1)(1 + p^n + \cdots + p^{n(m-1)}),$$

(1.9) and (1.10) imply

$$G'_{p^{nm}-1}(t) = \prod_{i=0}^{m-1} (\psi_i^{p^n-1} - F_i^{p^n-1}),$$

and it is easily verified that this becomes[3]

$$(1.11)\qquad G_{p^{nm}-1}(t) = \frac{\psi_m(t)}{t}.$$

Somewhat more generally, we may prove in the same way that

$$(1.12)\qquad G'_{p^{nm}-p^{ni}}(t) = \frac{\psi_m(t)}{\psi_i(t)} \qquad (0 \leqq i \leqq m).$$

From the definition (1.7) it is clear that we may write

$$(1.13)\qquad G_k(t) = t^k + \gamma_1^k t^{k-1} + \cdots + \gamma_k^k,$$

and that γ_i^k is integral, that is, a polynomial in x. Similarly

$$(1.14)\qquad G'_k(t) = t^k + \gamma_1'^k t^{k-1} + \cdots + \gamma_k'^k,$$

[2] This Journal, *An analogue of the von Staudt-Clausen theorem*, vol. 3(1937), pp. 503–517, p. 504.

[3] I, p. 141, formula (2.14).

where $\gamma_i'^k$ is integral. Note in particular that $\gamma_k^k = 0$ for $k > 0$, while $\gamma_k'^k \neq 0$ for

$$k = (p^n - 1)(p^{ne_1} + \cdots + p^{ne_s}) \qquad (e_1 < e_2 < \cdots < e_s),$$

indeed in this case

$$\gamma_k'^k = (-1)^s (F_{e_1} F_{e_2} \cdots F_{e_s})^{p^n - 1}. \tag{1.15}$$

In particular for $k = p^{nm} - 1$,

$$\gamma_k'^k = (-1)^m (F_1 \cdots F_{m-1})^{p^n - 1} = (-1)^m \frac{F_m}{L_m}; \tag{1.16}$$

the last equality follows from (1.4).

2. A polynomial of the form $f(t) = \sum \alpha_i t^{p^{ni}}$ may be called linear in that it has the properties

$$\begin{aligned} f(t + u) &= f(t) + f(u), \\ f(ct) &= cf(t) \qquad \text{for } c \text{ in } GF(p^n). \end{aligned} \tag{2.1}$$

In particular $\psi_m(t)$ is linear and satisfies (2.1). This is no longer true of $G_k(t)$. Using the notation (1.5), we have

$$G_k(ct) = \prod_{i=0}^{s} \psi_i^{\alpha_i}(ct) = \prod_{i=0}^{s} c^{\alpha_i} \psi_i^{\alpha_i}(t).$$

Since

$$\sum \alpha_i \equiv \sum \alpha_i p^{ni} \qquad (\mathrm{mod}\ p^n - 1),$$

it follows that

$$G_k(ct) = c^k G(t), \tag{2.2}$$

a result generalizing the second of (2.1).

As for the first of (2.1), note that

$$G_{\alpha p^{ni}}(t + u) = \{\psi_i(t) + \psi_i(u)\}^{\alpha} = \sum_{\beta=0}^{\alpha} \binom{\alpha}{\beta} \psi_i^{\beta}(t) \psi_i^{\alpha-\beta}(u),$$

and therefore

$$\begin{aligned} G_k(t + u) &= \prod_{i=0}^{s} \{\psi_i(t) + \psi_i(u)\}^{\alpha_i} \\ &= \prod_{i=0}^{s} \sum_{\beta_i=0}^{\alpha_i} \binom{\alpha_i}{\beta_i} \psi_i^{\alpha_i}(t) \psi_i^{\alpha_i - \beta_i}(u) \\ &= \sum_{\beta_i=0}^{\alpha_i} \binom{\alpha_0}{\beta_0} \cdots \binom{\alpha_s}{\beta_s} \psi_0^{\beta_0}(t) \cdots \psi_s^{\beta_s}(t) \psi_0^{\alpha_0 - \beta_0}(u) \cdots \psi_s^{\alpha_s - \beta_s}(u) \\ &= \sum_{\beta=0}^{k} \binom{k}{\beta} G_\beta(t) G_{k-\beta}(u), \end{aligned} \tag{2.3}$$

by (1.7) and a well-known property of binomial coefficients. For $k = p^{nm}$, (2.3) reduces to the first equation in (2.1). Another special case of interest is $k = p^{nm} - 1$. Replacing u by $-u$ we get

$$\begin{aligned} G_{p^{nm}-1}(t-u) &= \sum_{\beta=0}^{p^{nm}-1} \binom{p^{nm}-1}{\beta} G_\beta(t) G_{p^{nm}-1-\beta}(-u) \\ &= \sum_{\alpha+\beta=p^{nm}-1} G_\alpha(t) G_\beta(u) \end{aligned} \tag{2.4}$$

by (2.2) and the fact that $\binom{p^{nm}-1}{\beta} = (-1)^\beta$.

Turn next to G'_k; it is clear that (2.2) will certainly hold if each $\alpha_i < p^n - 1$. If some $\alpha_i = p^n - 1$, note that

$$G'_k(ct) = \psi_i^{p^n-1}(ct) - F_i^{p^n-1} = G'_k(t) \qquad (k = (p^n-1)p^{ni})$$

so that (2.2) will hold provided $c \neq 0$. Therefore we may assert that for all k

$$G'_k(ct) = c^k G'_k(t) \qquad (c \neq 0). \tag{2.5}$$

As for the analogue of (2.3), we shall prove the formula

$$\begin{aligned} G'_k(t+u) &= \sum_{\beta=0}^{k} \binom{k}{\beta} G'_\beta(t) G_{k-\beta}(u) \\ &= \sum_{\beta=0}^{k} \binom{k}{\beta} G_\beta(t) G'_{k-\beta}(u). \end{aligned} \tag{2.6}$$

If every $\alpha_i < p^n - 1$, there is nothing to prove. If some $\alpha_i = p^n - 1$, note that

$$\begin{aligned} G'_k(t+u) &= \{\psi_i(t) + \psi_i(u)\}^{p^n-1} - F_i^{p^n-1} \qquad (k = (p^n-1)p^{ni}) \\ &= \sum_{\alpha+\beta=p^n-1} \binom{p^n-1}{\alpha} \psi_i^\alpha(t)\psi_i^\beta(u) - F_i^{p^n-1} \\ &= \sum_{\alpha+\beta=p^n-1} \binom{p^n-1}{\alpha} G'_{\alpha p^{ni}}(t) G_{\beta p^{ni}}(u), \end{aligned}$$

and the first equality in (2.6) follows at once. The second equality is proved in exactly the same way.

As a special case of (2.6) we note

$$\begin{aligned} G'_{p^{nm}-1}(t-u) &= \sum_{\alpha+\beta=p^{nm}-1} G'_\alpha(t) G_\beta(u) \\ &= \sum_{\alpha+\beta=p^{nm}-1} G_\alpha(t) G'_\beta(u). \end{aligned} \tag{2.7}$$

To sum up we may state

THEOREM 1. *The polynomial $G_k(t)$ has the properties*

$$G_k(ct) = c^k G_k(t) \qquad (c \text{ in } GF(p^n)),$$

$$G_k(t + u) = (G(t) + G(u))^k,$$

where the right member is an abbreviation for (2.3); *the polynomial $G'_k(t)$ has the properties*

$$G'_k(ct) = c^k G'_k(t) \qquad (c \text{ in } GF(p^n),\ c \neq 0),$$

$$G'_k(t + u) = (G'(t) + G(u))^k = (G(t) + G'(u))^k.$$

3. In view of (1.12) it is evident that if $f(t)$ is any polynomial in t of degree k, we may put

$$f(t) = A_0 + A_1 G_1(t) + \cdots + A_k G_k(t), \tag{3.1}$$

where A_i is independent of t. The representation (3.1) is clearly unique. More generally, we may write

$$f(tu) = \sum_{i=0}^{k} A_i(u) G_i(t), \tag{3.2}$$

where $A_i(u)$ is a polynomial in u. However, (3.2) is actually no more general than (3.1), since we may take $g(t) = f(tu)$. Accordingly, we now limit ourselves to (3.1); the expansion (3.2) will be used in certain applications.

In (3.1) put $t = 0$ and we have $A_0 = f(0)$. We shall now determine the general coefficient in (3.1). Consider the polynomial

$$\Phi(t) = \sum_{\deg M < m} f(M) \frac{\psi_m(t)}{t - M}, \tag{3.3}$$

where $p^{nm} > k = \deg f(t)$. The degree of $\Phi(t)$ is at most $p^{nm} - 1$. If now M is any polynomial in x of degree $< m$, it is easily seen that (3.3) implies

$$\Phi(M) = f(M) \frac{\psi_m(t)}{t - M}\bigg]_{t=M}.$$

But

$$\frac{\psi_m(t)}{t - M}\bigg]_{t=M} = \frac{\psi_m(t - M)}{t - M}\bigg]_{t=M} = \frac{\psi_m(t)}{t}\bigg]_{t=0} = (-1)^m \frac{F_m}{L_m},$$

by (1.2) and (1.3). Therefore

$$\Phi(M) = (-1)^m \frac{F_m}{L_m} f(M). \tag{3.4}$$

Since (3.4) holds for the p^{nm} polynomials M of degree $< m$, and $p^{nm} > \deg f(t)$, it follows that

$$\Phi(t) = (-1)^m \frac{F_m}{L_m} f(t). \tag{3.5}$$

Making use of (3.3) and (3.5), we may compute the coefficients in (3.1). Indeed from (1.11) and (2.7) follows

$$\frac{\psi_m(t)}{t-M} = \frac{\psi_m(t-M)}{t-M} = G'_{p^{nm}-1}(t-M) = \sum_{i+j=p^{nm}-1} G_i(t)G'_j(M),$$

so that

$$\Phi(t) = \sum_{i+j=p^{nm}-1} G_i(t) \sum_{\deg M<m} G'_j(M)f(M).$$

If now we compare this with (3.2), we have at once (for $p^{nm} < k$)

$$(-1)^m \frac{F_m}{L_m} A_i = \sum_{\deg M<m} G'_{p^{nm}-1-i}(M)f(M). \tag{3.6}$$

However, a more general formula may be derived without any difficulty. Returning to (3.3) let m be arbitrary. Then (3.4) is still true but this will no longer imply (3.5). Let us break the right member of (3.2) into two parts

$$\sum_{i<p^{nm}} A_iG_i(t) + \sum_{i \geqq p^{nm}} A_iG_i(t) = f_1(t) + f_2(t), \tag{3.7}$$

say. Then the degree of $f_1(t)$ is at most $p^{nm} - 1$. Furthermore by (1.1) and (1.7)

$$G_i(M) = 0 \quad \text{for } i \geqq p^{nm}, \qquad \deg M < m,$$

and therefore $f_2(M) = 0$, or what amounts to the same thing,

$$f_1(M) = f(M) \quad \text{for } \deg M < m.$$

Since $\deg f_1(t) < p^{nm}$, we now get in place of (3.5)

$$\Phi(t) = (-1)^m \frac{F_m}{L_m} f_1(t) = (-1)^m \frac{F_m}{L_m} \sum_{i<p^{nm}} A_iG_i(t).$$

Hence repeating the argument that led to (3.6), we have the following

THEOREM 2. *Let $f(t)$ be a polynomial in t of degree $\leqq k$; then we have the unique representation*

$$f(t) = \sum_{i=0}^{k} A_iG_i(t).$$

Let $p^{nm} > i$. Then the coefficient A_i is determined by

$$(-1)^m \frac{F_m}{L_m} A_i = \sum_{\deg M<m} G'_{p^{nm}-1-i}(M)f(M). \tag{3.8}$$

It is not difficult to see directly that (3.6) and (3.8) are equivalent. Let $p^{ns} > p^{nm} > i$, then by (1.9) and (1.12),

$$G'_{p^{ns}-1-i} = G'_{p^{ns}-p^{nm}} G'_{p^{nm}-1-i} = \frac{\psi_s}{\psi_m} G'_{p^{nm}-1-i}.$$

Now since $\psi_s(M) = 0$ for deg $M < s$, $\psi_s(M)/\psi_m(M)$ will vanish unless deg $M < m$; in that case, we have

$$\left.\frac{\psi_s(t)}{\psi_m(t)}\right]_{t=M} = \left.\frac{\psi_s(t-M)}{\psi_m(t-M)}\right]_{t=M} = \left.\frac{\psi_s(t)}{\psi_m(t)}\right]_{t=0} = \frac{(-1)^s \dfrac{F_s}{L_s}}{(-1)^m \dfrac{F_m}{L_m}}.$$

Hence

$$A_i = (-1)^s \frac{L_s}{F_s} \sum_{\deg M<s} G'_{p^{ns}-1-i}(M) f(M) = (-1)^m \frac{L_m}{F_m} \sum_{\deg M<m} G'_{p^{nm}-1-i}(M) f(M),$$

and this shows that (3.8) remains true when m is increased. In particular, if s is so chosen that $p^{ns} > \deg f(t)$, we get (3.8).

In place of the polynomial $\Phi(t)$ defined by (3.3) we may define another polynomial

$$\Phi_1(t) = \sum_{\deg M=m}' f(M) \frac{\psi_m(t) - F_m}{t-M}, \tag{3.9}$$

the prime[4] on $\sum$ indicating that the summation is extended over primary M only of degree m. Then $\Phi_1(t)$ is of degree $< p^{nm}$. Then exactly as in the proof of (3.4)

$$\Phi_1(M) = (-1)^m \frac{F_m}{L_m} f(M),$$

for M primary of degree m. If now we take $p^{nm} > k = \deg f(t)$, we have at once

$$\Phi_1(t) = (-1)^m \frac{F_m}{L_m} f(t) = \Phi(t).$$

On the other hand, by (3.9)

$$\begin{aligned}\Phi_1(t) &= \sum_{\deg M=m}' f(M) \frac{\psi_m(t-M)}{t-M} \\ &= \sum_{\alpha+\beta=p^{nm}-1} G_\alpha(t) \sum_{\deg M=m}' f(M) G'_{p^{nm}-1-\alpha}(M).\end{aligned}$$

This leads to the following supplement to Theorem 2:

THEOREM 3. *With the notation of Theorem 2, the coefficient A_i is also given by*

$$(-1)^m \frac{F_m}{L_m} A_i = \sum_{\deg M=m}' G'_{p^{nm}-1-i}(M) f(M), \tag{3.10}$$

[4] We shall use this convention throughout.

where $p^{nm} >$ deg $f(t)$, *and the summation in the right member is over primary* M *only.*

Returning again to (3.3), we see clearly that (3.4) holds for arbitrary functions $f(t)$; however, (3.5) cannot in general be asserted. Let us assume that $f(t)$ admits the expansion

$$f(t) = \sum_{i=0}^{\infty} A_i G_i(t). \tag{3.11}$$

Break the right member of (3.11) into two parts

$$\sum_{i<p^{nm}} + \sum_{i \geq p^{nm}} = f_1(t) + f_2(t),$$

say, where m is arbitrary. Then as above $f_2(M) = 0$ for deg $M < m$, so that $f_1(M) = f(M)$. Also deg $f_1(t) < p^{nm}$. Since (3.4) holds and deg $\Phi(t) < p^{nm}$, we have

$$(-1)^m \frac{L_m}{F_m} \Phi(t) = f_1(t) = \sum_{i<p^{nm}} A_i G_i(t).$$

From this follows

THEOREM 4. *For* $m \geqq 0$, *and arbitrary* $f(t)$, *define the Lagrange polynomial*

$$\Phi(t) = \Phi(t; m, f) = \sum_{\deg M<m} f(M) \frac{\psi_m(t)}{t - M}, \tag{3.12}$$

so that

$$\Phi(M) = (-1)^m \frac{F_m}{L_m} f(M) \qquad (\deg M < m). \tag{3.13}$$

Then if $f(t)$ *admits the expansion* (3.11), *the coefficient* A_i *is given by* (3.8).

If (3.11) does not hold, we may still compute A_i by means of (3.8) and construct the polynomial

$$f_1(t) = \sum_{i<p^{nm}} A_i G_i(t),$$

so that $f_1(M) = f(M)$ for deg $M < m$. To indicate this situation we may write in place of (3.11)

$$f(t) \sim \sum A_i G_i(t), \tag{3.14}$$

which denotes merely that for A_i as given by (3.8), then

$$f(M) = \sum_{i<p^{nm}} A_i G_i(M) \qquad (\deg M < m).$$

Certain of the formulas of this section—namely, those that determine A_i—simplify considerably for special values of the subscript. For example, if in (3.8) we take $i = p^{nm} - 1$, we have

$$(-1)^m \frac{F_m}{L_m} A_{p^{nm}-1} = \sum_{\deg M<m} f(M). \tag{3.15}$$

Again in (3.10) take $i = p^{ns} - 1$, then by (1.12)

$$(-1)^m \frac{F_m}{L_m} A_{p^{ns}-1} = \sum_{\deg M=m}' f(M) \frac{\psi_m(M)}{\psi_s(M)}.$$

Since $\psi_m(M) = F_m$, this formula becomes

$$\frac{(-1)^m}{L_m} A_{p^{ns}-1} = \sum_{\deg M=m}' \frac{f(M)}{\psi_s(M)} \qquad (p^{nm} > k). \tag{3.16}$$

In particular for $s = 0$,

$$\frac{(-1)^m}{L_m} A_0 = \sum_{\deg M=m}' \frac{f(M)}{M}.$$

Similarly for $i = p^{ns}$, (3.10) yields

$$\frac{(-1)^m}{L_m} A_{p^{ns}} = \sum_{\deg M=m}' \frac{f(M)}{M} \frac{\psi_s^{p^n-1}(M)}{\psi_{s+1}(M)}.$$

4. Parallel to (3.2) we may consider the expansion of a polynomial $f(t)$ in terms of $G_i'(t)$:

$$f(t) = \sum_{i=0}^{k} A_i' G_i'(t). \tag{4.1}$$

We define $\Phi(t)$ as in (3.3), and prove without difficulty

THEOREM 5. *Let $f(t)$ be a polynomial of degree $\leqq k$, so that* (4.1) *holds. Then for $p^{nm} > k$, A_i' is determined by*

$$(-1)^m \frac{F_m}{L_m} A_i' = \sum_{\deg M<m} G_{p^{nm}-1-i}(M) f(M). \tag{4.2}$$

Again we may define $\Phi_1(t)$ as in (3.9), and derive the analogue of Theorem 3:

THEOREM 6. *With the notation of Theorem* 5, *we have*

$$(-1)^m \frac{F_m}{L_m} A_i' = \sum_{\deg M=m}' G_{p^{nm}-1-i}(M) f(M). \tag{4.3}$$

Note that Theorem 5 tells us somewhat less than the corresponding Theorem 2 of the previous section.

5. We now consider some special cases of the preceding formulas. Take first $f(t) = t^k$, then we write (3.1) in the form

$$t^k = \sum_{i=0}^{k} \mathfrak{A}_i^k G_i(t) \qquad (k \geqq 0), \tag{5.1}$$

so that $\mathfrak{A}_i^k$ is a polynomial in x. By Theorem 2 we have

$$(5.2)\qquad (-1)^m \frac{F_m}{L_m} \mathfrak{A}_i^k = \sum_{\deg M<m} M^k G'_{p^{nm}-1-i}(M) \qquad (p^{nm} > i),$$

while by Theorem 3 we have

$$(5.3)\qquad (-1)^m \frac{F_m}{L_m} \mathfrak{A}_i^k = \sum_{\deg M=m}' M^k G'_{p^{nm}-1-i}(M) \qquad (p^{nm} > k).$$

In particular by (3.15)

$$(5.4)\qquad (-1)^m \frac{F_m}{L_m} \mathfrak{A}_{p^{nm}-1}^k = \sum_{\deg M=m}' M^k,$$

and by (3.16)

$$(5.5)\qquad \frac{(-1)^m}{L_m} \mathfrak{A}_{p^{ns}-1}^k = \sum_{\deg M=m}' \frac{M^k}{\psi_s(M)} \qquad (p^{nm} > k).$$

If next we take $f(t) = t^k$ in (4.1), we may put

$$(5.6)\qquad t^k = \sum_{i=0}^k \mathfrak{A}_i'^k G_i'(t).$$

Then (4.2) implies

$$(5.7)\qquad (-1)^m \frac{F_m}{L_m} \mathfrak{A}_i'^k = \sum_{\deg M<m} M^k G_{p^{nm}-1-i}(M) \qquad (p^{nm} > k),$$

while (4.3) yields

$$(5.8)\qquad (-1)^m \frac{F_m}{L_m} \mathfrak{A}_i'^k = \sum_{\deg M=m}' M^k G_{p^{nm}-1-i}(M) \qquad (p^{nm} > k).$$

Formulas of a different kind involving $\mathfrak{A}_i^k$ and $\mathfrak{A}_i'^k$ will be found below (§7).

Consider now the case $f(t) = G_k(tu)$, an instance of (3.2). We write

$$(5.9)\qquad G_k(tu) = \sum_{i=0}^k \beta_i(u) G_i(t),$$

where $\beta_i(u) = \beta_i^k(u)$ is a polynomial in u. Then (3.8) and (3.10) imply

$$(5.10)\qquad (-1)^m \frac{F_m}{L_m} \beta_i(u) = \sum_{\deg M<m} G'_{p^{nm}-1-i}(M) G_k(Mu),$$

and

$$(5.11)\qquad (-1)^m \frac{F_m}{L_m} \beta_i(u) = \sum_{\deg M=m}' G_{p^{nm}-1-i}(M) G_k(Mu) \qquad (p^{nm} > k),$$

respectively.

In (5.9) put $u = 1$. Then clearly

$$\beta_i(1) = \beta_i^k(1) = \begin{cases} 0 & \text{for } i < k, \\ 1 & \text{for } i = k \end{cases}$$

Comparison with (5.10) leads to

$$\sum_{\deg M<m} G'_{p^{nm}-1-i}(M)G_k(M) = \begin{cases} 0 & \text{for } i < k, \\ (-1)^m \dfrac{F_m}{L_m} & \text{for } i = k, \end{cases}$$

while (5.11) yields

$$\sum_{\deg M=m}{}' G_{p^{nm}-1-i}(M)G_k(M) = \begin{cases} 0 & \text{for } i < k, \\ (-1)^m \dfrac{F_m}{L_m} & \text{for } i = k, \end{cases}$$

provided $p^{nm} > k$.

Making a slight change in notation, we may state

THEOREM 7. *For $l < p^{nm}$, k arbitrary,*

$$\text{(5.12)} \qquad \sum_{\deg M<m} G'_l(M)G_k(M) = \begin{cases} 0 & \textit{for } k + l \neq p^{nm} - 1, \\ (-1)^m \dfrac{F_m}{L_m} & \textit{for } k + l = p^{nm} - 1. \end{cases}$$

For $k < p^{nm}$, $l < p^{nm}$,

$$\text{(5.13)} \qquad \sum_{\deg M=m}{}' G'_l(M)G_k(M) = \begin{cases} 0 & \textit{for } k + l \neq p^{nm} - 1, \\ (-1)^m \dfrac{F_m}{L_m} & \textit{for } k + l = p^{nm} - 1. \end{cases}$$

In (5.12) the condition $l < p^{nm}$ may be removed provided the right member is modified slightly. Let $l \geqq p^{nm}$, so that

$$l = e + f = e_0p^{nm} + f, \qquad e_0 > 0, \qquad p^{nm} > f \geqq 0.$$

Then by (1.9) $G'_l(t) = G'_e(t)G'_f(t)$. But it is easily seen that for $\deg M < m$, $G'_e(M) = C$, a constant independent of M. Indeed

$$C = G(0) = 0$$

unless

$$\text{(5.14)} \qquad e = (p^n - 1)(p^{ne_1} + \cdots + p^{ne_s}), \qquad m \leqq e_1 < \cdots < e_s\,;$$

while if (5.14) holds then by (1.15) and (1.6),

$$C = (-1)^s(F_{e_1} \cdots F_{e_s})^{p^n-1} = (-1)^s g_e\,.$$

Thus the left member of (5.12) reduces to

$$C \sum_{\deg M<m} G'_f(M)G_k(M).$$

Since $f < p^{nm}$, (5.12) applies. Finally we see that for $k < p^{nm}$, the left member of (5.12) is 0 unless $f + k = p^{nm} - 1$ and (5.14) holds, in which case the sum

$$= (-1)^s g_e \cdot (-1)^m \frac{F_m}{L_m} = (-1)^{s+m} g_{k+l}\,.$$

Returning to (5.9) we remark that there are a number of similar formulas—namely, expansion of $G_k(tu)$ in terms of $G_i'(t)$, and $G_k'(tu)$ in terms of $G_i(t)$, $G_i'(t)$. For each the coefficients are determined by means of formulas similar to (5.10) and (5.11). For brevity this set of formulas will be omitted.

6. For the linear polynomial $\psi_s(t)$ there is the formula[5]

$$\sum_{\deg M=m}' \frac{\psi_s(M)}{M} = \begin{cases} (-1)^m \dfrac{F_m}{L_m} & \text{for } s = m, \\ 0 & \text{for } s \neq m. \end{cases} \tag{6.1}$$

We now prove the more general

$$\sum_{\deg M=m}' \frac{G_k(M)}{M} = \begin{cases} (-1)^m \dfrac{F_m^i}{L_m} & \text{for } k = ip^{nm},\ 0 \leqq i < p^n, \\ 0 & \text{otherwise.} \end{cases} \tag{6.2}$$

For $k = p^{ns}$, (6.2) reduces to (6.1).

The first half of (6.2) follows at once from[6]

$$\sum_{\deg M=m}' \frac{1}{M} = \frac{(-1)^m}{L_m}.$$

As for the second half, the case $0 < k < p^{nm}$ is an immediate consequence of the formula[7]

$$\sum_{\deg M=m}' M^{k-1} = 0 \qquad (0 < k < p^{nm}).$$

Suppose next $k = ip^{nm} + j$, $0 < i < p^n$, $0 < j < p^{nm}$; then by (1.7) $G_k = \psi_m^i G_j$ and therefore

$$\sum_M' \frac{G_k(M)}{M} = F_m^i \sum_M' \frac{G_j(M)}{M} = 0,$$

by the preceding case. Finally let $k \geqq p^{n(m+1)}$, so that $G_k(M) = 0$ and (6.2) is surely satisfied. This completes the proof of the formula.

As an application of (6.2) let $f(t)$ be an arbitrary polynomial of degree k, and as in (3.1) put

$$f(t) = \sum_{i=0}^{k} A_i G_i(t). \tag{6.3}$$

Then

$$\sum_{\deg M=m}' \frac{f(M)}{M} = \sum_{i=0}^{k} A_i \sum_M' \frac{G_i(M)}{M}.$$

[5] This Journal, *Some sums involving polynomials in a Galois field*, vol. 5(1939), pp. 941–947, p. 943.

[6] I, p. 160, formula (9.09).

[7] This follows from I, Theorem 9.5.

If now we apply (6.2) we get

$$\sum_{\deg M=m}' \frac{f(M)}{M} = \frac{(-1)^m}{L_m} \sum_{i=0}^{p^n-1} A_{ip^{nm}} F_m^i. \tag{6.4}$$

For further simplification, let $f(0) = 0$, so that $A_0 = 0$. Also assume that $f(t)$ satisfies

$$f(ct) = c^\lambda f(t) \qquad (c \text{ in } GF(p^n)), \tag{6.5}$$

where λ is fixed; we may suppose that $0 < \lambda \leqq p^n - 1$. Now by (2.2), $G_i(ct) = c^i G_i(t)$. Combining this with (6.5), we have at once

$$A_i = 0 \quad \text{for } i \not\equiv \lambda \pmod{p^n - 1},$$

so that only one term on the right of (6.4) remains:

$$\sum_{\deg M=m}' \frac{f(M)}{M} = (-1)^m \frac{F_m^\lambda}{L_m} A_{\lambda p^{nm}}. \tag{6.6}$$

The condition (6.5) will be satisfied for $f(t) = t^k$, $k \geqq 1$, and

$$k \equiv \lambda \pmod{p^n - 1} \qquad (0 < \lambda \leqq p^n - 1).$$

Then in the notation of (5.1), we get from (6.6):

$$\sum_{\deg M=m}' M^{k-1} = (-1)^m \frac{F_m}{L_m} \mathfrak{A}^k_{\lambda p^{nm}}. \tag{6.7}$$

In particular for $k \equiv 1 \pmod{p^n - 1}$ this becomes

$$\sum_{\deg M=m}' M^{k-1} = (-1)^m \frac{F_m}{L_m} \mathfrak{A}^k_{p^{nm}}.$$

7. Returning to §5 we now derive some additional formulas involving the coefficients $\mathfrak{A}_i^k$. From (5.1) we get

$$\begin{aligned} \sum_{k=0}^{p^{nm}-1} \frac{t^k}{u^{k+1}} &= \sum_{k=0}^{p^{nm}-1} \frac{1}{u^{k+1}} \sum_{i=0}^{k} \mathfrak{A}_i^k G_i(t) \\ &= \sum_{i=0}^{p^{nm}-1} G_i(t) \sum_{k=i}^{p^{nm}-1} \frac{\mathfrak{A}_i^k}{u^{k+1}}. \end{aligned} \tag{7.1}$$

Now the left member

$$= \frac{1}{u}\left(1 - \frac{t}{u}\right)^{p^{nm}-1} = \frac{(u-t)^{p^{nm}-1}}{u^{p^{nm}}},$$

so that (7.1) becomes

$$(u - t)^{p^{nm}-1} = \sum_{i=0}^{p^{nm}-1} G_i(t) \sum_{k=i}^{p^{nm}-1} \mathfrak{A}_i^k u^{p^{nm}-1-k}. \tag{7.2}$$

Again we have the identity[8]

$$t^{p^{nm}} = \sum_{i=0}^{m} \begin{Bmatrix} m \\ i \end{Bmatrix} \psi_i(t),$$

where for brevity we put

$$\begin{Bmatrix} m \\ i \end{Bmatrix} = \frac{F_m}{F_i F_{m-i}^{p^{ni}}},$$

which by (1.11) becomes

$$t^{p^{nm}-1} = \sum_{i=0}^{m} \begin{Bmatrix} m \\ i \end{Bmatrix} G'_{p^{ni}-1}(t),$$

and by (2.7) this implies

$$\begin{aligned}(u - t)^{p^{nm}-1} &= \sum_{k=0}^{m} \begin{Bmatrix} m \\ k \end{Bmatrix} \sum_{i+j=p^{nk}-1} G'_j(u) G_i(t) \\ &= \sum_{i=0}^{p^{nm}-1} G_i(t) \sum_{p^{nk}>i} \begin{Bmatrix} m \\ k \end{Bmatrix} G'_{p^{nk}-1-i}(u).\end{aligned}$$

Comparison with (7.2) leads to

$$\sum_{k=i}^{p^{nm}-1} \mathfrak{A}_i^k u^{p^{nm}-1-k} = \sum_{p^{nk}>i} \begin{Bmatrix} m \\ k \end{Bmatrix} G'_{p^{nk}-1-i}(u). \tag{7.3}$$

Let $p^{ns} > i \geqq p^{n(s-1)}$. Then since

$$G'_{p^{nk}-1-i} = G'_{p^{nk}-p^{ns}} G'_{p^{ns}-1-i} = \frac{\psi_k}{\psi_s} G'_{p^{ns}-1-i},$$

we may rewrite (7.3) in the form

$$\sum_{k \geqq i} \mathfrak{A}_i^k u^{p^{nm}-1-k} = \frac{G'_{p^{ns}-1-i}(u)}{\psi_s(u)} \sum_{k=s}^{m} \begin{Bmatrix} m \\ k \end{Bmatrix} \psi_k(u). \tag{7.4}$$

In particular for $i = p^{ns} - 1$, this reduces to

$$\sum \mathfrak{A}_{p^{ns}-1}^k u^{p^{nm}-1-k} = \frac{1}{\psi_s(u)} \sum_{k=s}^{m} \begin{Bmatrix} m \\ k \end{Bmatrix} \psi_k(u).$$

In exactly the same way we may derive an identity similar to (7.3) involving the $\mathfrak{A}_i'^k$ of (5.6). The final formula is

$$\begin{aligned}\sum_{k=i}^{p^{nm}-1} \mathfrak{A}_i'^k u^{p^{nm}-1-k} &= \sum_{p^{nk}>i} \begin{Bmatrix} m \\ k \end{Bmatrix} G_{p^{nk}-1-i}(u) \\ &= G_{p^{ns}-1-i}(u) \sum_{k=s}^{m} \begin{Bmatrix} m \\ k \end{Bmatrix} (\psi_s \psi_{s+1} \cdots \psi_{k-1})^{p^n-1}.\end{aligned} \tag{7.5}$$

[8] I, p. 144, formula (3.09).

In the next place by (1.2) we have

$$G'_{p^{nm}-1}(u-t) = \frac{\psi_m(u-t)}{u-t} = \sum_{k=0}^{m} (-1)^{m-k} \begin{bmatrix} m \\ k \end{bmatrix} (u-t)^{p^{nk}-1}. \tag{7.6}$$

But

$$\begin{aligned}(u-t)^{p^{nk}-1} &= \sum_{j=0}^{p^{nk}-1} u^{p^{nk}-1-j} t^j = \sum_{j=0}^{p^{nk}-1} u^{p^{nk}-1-j} \sum_{i=0}^{j} \mathfrak{A}_i^j G_i(t) \\ &= \sum_{i=0}^{p^{nk}-1} G_i(t) \sum_{j=i}^{p^{nk}-1} \mathfrak{A}_i^j u^{p^{nk}-1-j}.\end{aligned}$$

Substituting in (7.5), we get

$$\begin{aligned}G'_{p^{nm}-1}(u-t) &= \sum_{i=0}^{p^{nm}-1} G_i(t) \sum_{p^{nk}>i} (-1)^{m-k} \begin{bmatrix} m \\ k \end{bmatrix} \sum_{j=i}^{p^{nk}-1} \mathfrak{A}_i^j u^{p^{nk}-1-j} \\ &= \sum_{i=0}^{p^{nm}-1} G_i(t) \sum_{s=0}^{p^{nm}-1-i} u^s \sum_{p^{nk}>i+s} (-1)^{m-k} \begin{bmatrix} m \\ k \end{bmatrix} \mathfrak{A}_i^{p^{nk}-1-s}.\end{aligned} \tag{7.7}$$

Therefore by comparing with (2.7), we have

$$G'_{p^{nm}-1-i}(u) = \sum_{s=0}^{p^{nm}-1-i} u^s \sum_{p^{nk}>i+s} (-1)^{m-k} \begin{bmatrix} m \\ k \end{bmatrix} \mathfrak{A}_i^{p^{nk}-1-s}, \tag{7.8}$$

which may also be written as

$$G'_i(u) = \sum_{s=0}^{i} u^s \sum_{k} (-1)^{m-k} \begin{bmatrix} m \\ k \end{bmatrix} \mathfrak{A}_{p^{nm}-1-i}^{p^{nk}-1-s}, \tag{7.9}$$

the inner sum extending over all $k \leqq m$, and $p^{nk} > p^{nm} - 1 - i + s$. From (7.9) we have at once a formula for $\gamma_i'^k$ of (1.14).

Analogous to (7.8) we have also

$$G_{p^{nm}-1-i}(u) = \sum_{s=0}^{i} u^s \sum_{p^{nk}>i+s} (-1)^{m-k} \begin{bmatrix} m \\ k \end{bmatrix} \mathfrak{A}_i'^{p^{nk}-1-s}; \tag{7.10}$$

there are also various formulas connecting γ_i^k and $\mathfrak{A}_i'^k$ that we shall not go into.

The question naturally arises of a simple generating function for $G_i(t)$. Consider the polynomial in t of degree $< p^{nm}$:

$$\Gamma(t) = \Gamma_m(t) = \frac{\psi_m(u-t)}{(u-t)\psi_m(u)}. \tag{7.11}$$

Evidently

$$\Gamma(M) = \frac{1}{u-M} \qquad (\deg M < m). \tag{7.12}$$

This suggests a close connection with Theorem 4. Indeed for $f(t) = 1/(u-t)$, (3.12) becomes

$$\Phi(t) = \Phi\left(t; m, \frac{1}{u-t}\right) = \sum_{\deg M<m} \frac{1}{u-M} \frac{\psi_m(t)}{t-M};$$

now clearly $\Phi(t)$ is a polynomial in t of degree $< p^{nm}$, and by (3.13)

$$\Phi(M) = (-1)^m \frac{F_m}{L_m} \frac{1}{u - M} \qquad (\deg M < m).$$

Comparison with (7.12) shows that

$$\Phi(t) = (-1)^m \frac{F_m}{L_m} \Gamma_m(t). \tag{7.13}$$

Now (3.11) does not hold in the present case. (The polynomial $\Gamma_m(t)$ does in fact approach a limit as $m \to \infty$, but the limit $\neq 1/(u - t)$.) Thus the discussion immediately following Theorem 4 applies here. We first compute the coefficients A_i of (3.11). Essentially this has been done at the beginning of this section. We have

$$\Gamma_m(t) = \frac{1}{\psi_m(u)} G'_{p^{nm}-1}(u - t) = \frac{1}{\psi_m(u)} \sum_{i+j=p^{nm}-1} G'_j(u)G_i(t), \tag{7.14}$$

so that

$$A_i = \frac{G'_{p^{nm}-1-i}(u)}{\psi_m(u)}. \tag{7.15}$$

Now by (7.14) the difference

$$\Gamma_m(t) - \Gamma_{m-1}(t) = \sum_{i=0}^{p^{n(m-1)}-1} G_i(t)\left\{\frac{G'_{p^{nm}-1-i}(u)}{\psi_m(u)} - \frac{G'_{p^{n(m-1)}-1-i}(u)}{\psi_{m-1}(u)}\right\} + \frac{1}{\psi_m(u)} \sum_{i=p^{n(m-1)}}^{p^{nm}-1} G_i(t)G_{p^{nm}-1-i}(u),$$

and the quantity in braces

$$= \frac{1}{\psi_m(u)} \{G'_{p^{nm}-1-i} - G'_{p^{nm}-p^{n(m-1)}}G'_{p^{n(m-1)}-1-i}\} = 0.$$

Thus (7.14) becomes

$$\Gamma_m(t) = \frac{1}{u} + \sum_{k=1}^{m} \frac{1}{\psi_k(u)} \sum_{i=p^{n(k-1)}}^{p^{nk}-1} G_i(t)G'_{p^{nk}-1-i}(u). \tag{7.16}$$

Hence in the notation of (3.14)

$$\frac{1}{u - t} \sim \frac{1}{u} + \sum_{i=1}^{\infty} G_i(t) \frac{G'_{p^{nk}-1-i}(u)}{\psi_k(u)}, \tag{7.17}$$

where k is determined by

$$p^{n(k-1)} \leqq i < p^{nk}.$$

It will be recalled that (7.17) means that the two members are equal for $t = M$, $\deg M < m$, m arbitrary, the sum in the right member extending over $i < p^{nm}$.

Since A_i may also be defined by (3.8) we have the identity

$$\sum_{\deg M<m} \frac{1}{u-M} G'_{p^{nm}-1-i}(M) = (-1)^m \frac{F_m}{L_m} \frac{G'_{p^{nk}-1-i}(u)}{\psi_k(u)}.$$

as may be checked directly.

8. We conclude with some arithmetical applications. Let

$$f(t) = \sum_{i=0}^{k} A_i G_i(t);$$

we shall call $f(t)$ integral-valued if $f(M)$ is integral—that is, a polynomial in x—for all polynomials M. We consider first the case of "linear" $f(t)$:

$$f(t) = \sum_{i=0}^{k} \alpha_i t^{p^{ni}} = \sum_{i=0}^{k} A_i \psi_i(t), \tag{8.1}$$

and prove the following

THEOREM 8. *A linear polynomial $f(t)$ is integral-valued if and only if $A_i F_i$ is integral, where A_i is defined by* (8.1).

We shall require

LEMMA 1. *For all M, $\psi_i(M)/F_i$ is integral.*

It suffices to prove $\psi_i(x^m)/F_i$ integral. This follows by induction from the identity[9]

$$\psi_i(x^{m+1}) = x\psi_i(x^m) + (x^{p^{nm}} - x)\psi_{i-1}^{p^n}(x^m),$$

which may be written

$$\frac{\psi_i(x^{m+1})}{F_i} = x\frac{\psi_i(x^m)}{F_i} + \left(\frac{\psi_{i-1}(x^m)}{F_{i-1}}\right)^{p^n}, \tag{8.2}$$

and the fact that $\psi_i(x^i) = F_i$.

In view of Lemma 1 it is clear that if $A_i F_i$ is integral, then $f(t)$ is integral-valued. It remains to prove the necessity of the condition. This may be done by a familiar method.[10] Since $\psi_i(x^m) = 0$ for $i > m$, it follows from (8.2) that

$$f(1) = A_0,$$
$$f(x) = A_0 x + A_1 F_1,$$
$$\cdots\cdots\cdots\cdots$$
$$f(x^m) = A_0 x^m + \cdots + A_{m-1}\psi_{m-1}(x^m) + A_m F_m,$$

so that we may compute $A_0, A_1F_1, \cdots, A_mF_m$. If we apply Lemma 1 again, it is clear that A_mF_m is integral.

[9] I, p. 141, formula (2.13).

[10] See, for example, Pólya-Szegö, *Aufgaben und Lehrsätze aus der Analysis*, vol. 2, 1925. p. 339.

For the general case we shall prove

THEOREM 9. *The polynomial*

$$f(t) = \sum_{i=0}^{k} A_i G_i(t) \tag{8.3}$$

is integral-valued if and only if $A_i g_i$ *is integral;* g_i *is defined by* (1.5) *and* (1.6).

We first extend Lemma 1.

LEMMA 2. *For all* M, $G_k(M)/g_k$ *and* $G_k'(M)/g_k$ *are integral.*

Indeed with the notation (1.5),

$$\frac{G_k(M)}{g_k} = \prod_{i=0}^{s} \left(\frac{\psi_i(M)}{F_i}\right)^{\alpha_i},$$

which is integral by Lemma 1; the proof for $G_k'(M)/g_k$ is similar.

If now in (8.3) $A_i g_i$ is integral, it follows at once from Lemma 2 that $f(t)$ is integral-valued. However, the necessity of the theorem cannot be proved in the same manner as before.[11] Instead we use (3.8), so that

$$(-1)^m g_{p^{nm}-1} A_i = \sum_{\deg M < m} G'_{p^{nm}-1-i}(M) f(M). \tag{8.4}$$

By Lemma 2, each term on the right is divisible by $g_{p^{nm}-1-i}$. Also by (1.6)

$$g_{p^{nm}-1} = g_{p^{nm}-1-i}\, g_i\,,$$

so that (8.4) may be written

$$(-1)^m g_i A_i = \sum_M \frac{G_{p^{nm}-1-i}(M)}{g_{p^{nm}-1-i}} f(M),$$

which is integral. This completes the proof of the theorem.

In precisely the same way we may prove

THEOREM 10. *The polynomial*

$$f(t) = \sum_{i=0}^{k} A_i' G_i'(t)$$

is integral-valued if and only if $A_i' g_i$ *is integral.*

The quantities appearing in (8.2) are of some interest.[12] Let

$$H_i^s = \frac{\psi_i(x^{i+s})}{F_i},$$

so that by Lemma 1, H_m^s is integral for $s \geqq 0$. In place of (8.2) we may write

$$H_i^{s+1} = xH_i^s + (H_{i-1}^{s+1})^{p^n}. \tag{8.5}$$

[11] The method employed in the proof of Theorem 8 may be used if first we introduce a certain set of auxiliary polynomials in t. However, the present proof is much shorter.

[12] See *A class of polynomials*, Transactions of the American Mathematical Society, vol. 43(1938), pp. 167–182.

Since $H_i^0 = 1$, we have

$$H_i^1 = x + (H_{i-1}^1)^{p^n}.$$

From this it follows that

$$H_i^1 = x + x^{p^n} + \cdots + x^{p^{ni}}.$$

Next for $s = 1$, (8.5) becomes

$$H_i^2 = xH_i^1 + (H_{i-1}^2)^{p^n},$$

and from this it follows that

$$H_i^2 = \sum_{0 \leq \alpha \leq \beta \leq i} x^{p^{n\alpha}} x^{p^{n\beta}}.$$

Proceeding in this way we may prove that

$$H_i^s = \sum_{0 \leq \alpha_1 \leq \cdots \leq \alpha_s \leq i} x^{p^{n\alpha_1}} \cdots x^{p^{n\alpha_s}},$$

the sum of the products of $x, x^{p^n}, \cdots, x^{p^{ni}}$ taken s at a time, repetitions allowed. Other properties of H_i^s will be considered later.

DUKE UNIVERSITY.

APPROXIMATION TO FUNCTIONS BY TRIGONOMETRIC POLYNOMIALS

By A. C. Offord

1. Let $f(x)$ be any integrable function. It is well known that there is just one trigonometric polynomial of order n which coincides with $f(x)$ in the $2n + 1$ points

$$x_i = i\frac{2\pi}{2n+1} \qquad (i = 0, 1, \cdots, 2n).$$

The explicit expression for these polynomials is very simple.[1] If we write

$$\varphi_n(t) = i\frac{2\pi}{n} \qquad (i = 0, 1, 2, \cdots, n-1) \tag{1}$$

for $2\pi i/n \leqq t < 2\pi\,(i + 1)/n$, the desired polynomials are

$$\begin{aligned} U_n(f, x) &= \frac{1}{2\pi}\int_0^{2\pi} f(t)\,\frac{\sin\,(n+\frac{1}{2})(x-t)}{\sin\,\frac{1}{2}(x-t)}\,d\varphi_{2n+1}(t) \\ &= \frac{1}{2n+1}\sum_{i=0}^{2n} f(x_i)\,\frac{\sin\,(n+\frac{1}{2})(x-x_i)}{\sin\,\frac{1}{2}(x-x_i)}, \end{aligned} \tag{2}$$

and plainly

$$U_n(f, x_i) = f(x_i) \qquad \text{for all } i.$$

In view of the close resemblance between (2) and Dirichlet's sum for the first n terms of the Fourier series of $f(x)$, we may expect to find some analogy between the behavior, for large n, of the polynomials U_n and the sums s_n of the Fourier series of $f(x)$. This is indeed the case. In some ways, however, interpolating polynomials have an advantage over partial sums. If the function is very smooth, polynomials give a much better approximation to the function than partial sums of the same order. Thus, if $f(x)$ is analytic, the respective errors after the n-th term are, as is well known, $(\frac{1}{2})^{n+1}4M/(n+1)!$ and $M/(n+1)!$ respectively, where M is the least upper bound of $|f^{(n+1)}(x)|$.

The advantage of the interpolating polynomials holds only for smooth functions. As the function becomes less and less smooth, the advantage decreases becoming ultimately a disadvantage. Thus Marcinkiewicz[2] and Grünwald[3]

Received September 15, 1939. My warmest thanks are due to Dr. Marcinkiewicz for some very interesting discussions from which I derived much help.

[1] See, e.g., E. Feldheim, *Théorie de la Convergence des Procédés d'Interpolation*, Mémorial des Sciences Mathématiques, no. 95, Paris, 1939, p. 22.

[2] *Interpolating polynomials for absolutely continuous functions* (in Polish), Wiadomosci Mat., vol. 39(1935), pp. 85–115; and *Sur la divergence des polynomes d'interpolation*, Acta Litterarum ac Scientiarum, vol. 8(1937), pp. 131–135.

[3] *Über Divergenzerscheinungen der Lagrangeschen Polynome*, Acta Litterarum ac Scientiarum, vol. 7(1935), pp. 207–221.

have shown that corresponding to any sequence $\{n_i\}$ there exists a continuous function for which $U_{n_i}(f, x)$ diverges for all x. On the other hand, it is well known that the partial sums $s_{n_i}(x)$ of the Fourier series of a continuous function converge for almost all x, whenever $n_{i+1}/n_i \geqq k > 1$.

Various attempts have been made to construct trigonometric polynomials which converge to $f(x)$ when $f(x)$ is continuous. The fact that there exist such polynomials, defined once for all, was proved by Borel.[4] Probably the most interesting solution was given by Jackson[5] who showed that the polynomials

$$J_n(f, x) = \frac{1}{2(n+1)\pi} \int_0^{2\pi} f(t) \left\{ \frac{\sin \frac{1}{2}(n+1)(x-t)}{\sin \frac{1}{2}(x-t)} \right\}^2 d\varphi_{n+1}(t)$$

$$= \frac{1}{(n+1)^2} \sum_{i=0}^{n} f(x_i) \left\{ \frac{\sin \frac{1}{2}(n+1)(x-x_i)}{\sin \frac{1}{2}(x-x_i)} \right\}^2$$

have the desired property. These polynomials have, however, two disadvantages. In the first place $J_n(f, x)$ is equal to $f(x)$ only at $n + 1$ points and to get as good agreement as in (2) we must take twice as many terms. Secondly, if $f(x)$ is not continuous, their behavior is no better than the $U_n(f, x)$.[6]

Marcinkiewicz[7] has introduced the following polynomials. He writes

$$U_{n,m}(f, x) = \frac{1}{2\pi} \int_0^{2\pi} f(t) \frac{\sin (m + \frac{1}{2})(x-t)}{\sin \frac{1}{2}(x-t)} d\varphi_{2n+1}(t),$$

$$\sigma_n(f, x) = \frac{1}{n+1} \sum_{m=0}^{n} U_{n,m}(f, x)$$

$$= \frac{1}{2\pi(n+1)} \int_0^{2\pi} f(t) \left\{ \frac{\sin (n+1)(x-t)}{\sin \frac{1}{2}(x-t)} \right\}^2 d\varphi_{2n+1}(t).$$

These polynomials certainly converge to $f(x)$ if $f(x)$ is continuous and they are not open to the objection of the Jackson polynomials in that they require so many more terms to give as good agreement. Of course, it is no longer true that $\sigma_n(f, x)$ coincides with $f(x)$ at the points x_i. Unfortunately, if $f(x)$ is not continuous they are no better than the ordinary polynomials.

[4] Cf. Feldheim, loc. cit. (see footnote 1), p. 25.

[5] *Theory of Approximation*, New York, 1930. Cf. also Feldheim, loc. cit. (see footnote 1), p. 26.

While the present paper was in press, there appeared an interesting paper by J. Favard, *Sur l'interpolation*, Bull. Soc. Math. France, vol. 67 (1939), pp. 102–113, in which the author gives a very simple proof of Borel's theorem in the general case for any continuous function.

[6] Cf. J. Marcinkiewicz and A. Zygmund, *Mean values of trigonometrical polynomials*, Fundamenta Math., vol. 28(1937), pp. 131–166; p. 163.

[7] *Sur l'interpolation*, Studia Math., vol. 6(1936), pp. 1–17.

2. The object of this note is first to define trigonometric polynomials which converge uniformly whenever $f(x)$ is continuous and secondly to show how this definition can be generalized so as to obtain a class of polynomials which converge almost everywhere for any integrable function $f(x)$.

Our first result, given in Theorem 1 below, furnishes another solution of Borel's problem. Borel's theorem has no strict analogue for general integrable functions as such a function is not determined by its values at an enumerable set of points. Theorem 2, however, provides us with a method for constructing convergent polynomials $\sum_{\nu=0}^{n} A_{n,\nu} P_{n,\nu}(x)$, where the $P_{n,\nu}(x)$ are fixed once for all and the $A_{n,\nu}$ are readily determined by the values of $f(x)$ in the immediate neighborhood of each point x_i.[8]

THEOREM 1. *If $f(x)$ is a continuous function and $I_n(f, x)$ denotes the trigonometric polynomial of order n which takes the values*

$$\tfrac{1}{2}f(x_i) + \tfrac{1}{4}\{f(x_{i+1}) + f(x_{i-1})\}$$

at the $2n + 1$ points

$$x_i = i\,\frac{2\pi}{2n+1} \qquad (i = 0, 1, \cdots, 2n),$$

then $I_n(f, x)$ converges uniformly to $f(x)$.

Proof. In view of (1) and (2),

$$U_n\left\{f\left(t + \frac{2\pi}{2n+1}\right), x\right\} = U_n\left\{f(t), x + \frac{2\pi}{2n+1}\right\},$$

and so

$$\begin{aligned} I_n(f, x) &= \frac{1}{2}U_n(f, x) + \frac{1}{4}U_n\left(f, x + \frac{2\pi}{2n+1}\right) + \frac{1}{4}U_n\left(f, x - \frac{2\pi}{2n+1}\right) \\ &= \frac{1}{2\pi}\int_0^{2\pi} f(t)K_n(x, t)\,d\varphi_{2n+1}(t), \end{aligned}$$

where

$$\begin{aligned} K_n(x, t) &= \frac{1}{2}\sin\,(n + \tfrac{1}{2})(x - t) \\ &\cdot\left\{\frac{1}{\sin\frac{1}{2}(x - t)} - \frac{1}{2\sin\frac{1}{2}\left(x - t + \frac{2\pi}{2n+1}\right)} - \frac{1}{2\sin\frac{1}{2}\left(x - t - \frac{2\pi}{2n+1}\right)}\right\}. \end{aligned}$$

[8] Our results resemble somewhat Rogosinski's theorems on Fourier series. Cf. W. Rogosinski, *Über die Abschnitte trigonometrischer Reihen*, Math. Annalen, vol. 95(1925), pp. 110–134, and *Reihensummierung durch Abschnittskoppelungen*, Math. Zeitschrift, vol. 25(1926), pp. 132–149. Cf. also A. Zygmund, *Trigonometric Series*, Warsaw, 1935, p. 181.

Rogosinski, however, averages the partial sums s_n and not the function. Averaging the interpolating polynomials $U_n(f, x)$ does appear to yield the same result. Cf. Marcinkiewicz and Zygmund, loc. cit. (see footnote 6), p. 164.

The following inequalities for $K_n(x, t)$ are immediate:

(3) $$|K_n(x, t)| \leqq Cn,$$

(4) $$|K_n(x, t)| \leqq \frac{C}{n(x-t)^2}, \qquad |x - t| \geqq \frac{1}{n};$$

whence

$$|I_n(f, x) - f(x)| \leqq \epsilon n \int_{x-1/n}^{x+1/n} d\varphi_{2n+1}(t) + \frac{\epsilon}{n} \int_{1/n}^{\delta} \frac{d\varphi_{2n+1}(t)}{t^2} + O\left(\frac{1}{n}\right),$$

and the desired result follows at once.

3. We proceed to the case in which $f(x)$ is any Lebesgue integrable function. Plainly, it is no longer sufficient to take into account only the values of $f(x)$ at an enumerable set of points, so this time we average the function over a small interval. Let us write

$$\Delta_n = \frac{2\kappa_n \pi}{2n + 1},$$

where κ_n is an integer which may depend on n but which is, at any rate, such that Δ_n tends to zero. The most interesting case is that in which $\kappa_n = 1$. Consider the function

(5) $$f_n(x) = \frac{1}{2\Delta_n} \int_{x-\Delta_n}^{x+\Delta_n} f(t)\, dt.$$

Since Δ_n tends to zero, $f_n(x)$ tends to $f(x)$ at all points of the Lebesgue set. If $f(x)$ is very smooth, $f_n(x)$ agrees closely with the function in Theorem 1.

THEOREM 2. *The interpolating polynomials $I_n(f, x)$ which take the values $f_n(x_i)$ at the $2n + 1$ points*

$$i \frac{2\pi}{2n+1} \qquad (i = 0, 1, 2, \cdots, 2n)$$

converge to $f(x)$ at all points of the Lebesgue set.

Proof. We have, by (2),

$$I_n(f, x) = U_n(f_n, x) = \frac{1}{2\pi} \int_0^{2\pi} f_n(t) \frac{\sin (n + \frac{1}{2})(x - t)}{\sin \frac{1}{2}(x - t)}\, d\varphi_{2n+1}(t).$$

Let

$$\Phi_n(t) = \int_0^t \frac{\sin (n + \frac{1}{2})(x - u)}{\sin \frac{1}{2}(x - u)}\, d\varphi_{2n+1}(u),$$

so that

$$I_n(f, x) = \frac{1}{2\pi} [f_n(t)\Phi_n(t)]_0^{2\pi} - \frac{1}{2\pi} \int_0^{2\pi} f_n'(t)\Phi_n(t)\, dt.$$

But $f_n(t)$ and $\Phi_n(t)$ are of period 2π and

$$f_n'(t) = \frac{1}{2\Delta_n}\{f(t+\Delta_n) - f(t-\Delta_n)\}.$$

Hence

$$I_n(f, x) = -\frac{1}{2\pi}\frac{1}{2\Delta_n}\int_0^{2\pi}\{f(t+\Delta_n) - f(t-\Delta_n)\}\Phi_n(t)\,dt$$
$$= \frac{1}{2\pi}\frac{1}{2\Delta_n}\int_0^{2\pi} f(t)K_n(x, t)\,dt,$$

where

$$K_n(x, t) = \frac{1}{2\Delta_n}\{\Phi_n(t+\Delta_n) - \Phi_n(t-\Delta_n)\}$$
$$= \frac{1}{2\Delta_n}\int_{t-\Delta_n}^{t+\Delta_n}\frac{\sin(n+\frac{1}{2})(x-u)}{\sin\frac{1}{2}(x-u)}\,d\varphi_{2n+1}(u).$$

We prove the following inequalities for $K_n(x, t)$:

(6) $$|K_n(x, t)| \leqq Cn;$$

(7) $$|K_n(x, t)| \leqq \frac{C}{n(x-t)^2}, \qquad |x - t| \geqq 3\Delta_n,$$

except possibly for a finite number of values of t;

(8) $$|K_n(x, t)| \leqq \frac{C}{\Delta_n} + \frac{C}{n(x-t)^2}, \qquad |x - t| \geqq \frac{3}{n},$$

except possibly for a finite number of values of t.

The first inequality is immediate. We proceed to the proof of (7). Write $\beta_k = 2k\pi/(2n+1)$ and suppose $t \neq \beta_k$. This excludes only a finite number of values of t. Let k_1 and k_2 be defined by

$$\beta_{k_1-1} < t - \Delta_n \leqq \beta_{k_1} < \beta_{k_2} \leqq t + \Delta_n < \beta_{k_2+1}.$$

Since Δ_n is of the form $2\kappa_n\pi/(2n+1)$, where κ_n is an integer, the interval $(t - \Delta_n, t + \Delta_n)$ must contain an even number of points β_k. Then

$$2\Delta_n K_n(x, t) = \frac{1}{2n+1}\sum_{k=k_1}^{k_2}\frac{\sin(n+\frac{1}{2})(x-\beta_k)}{\sin\frac{1}{2}(x-\beta_k)}$$
$$= \frac{\sin(n+\frac{1}{2})x}{2n+1}\sum_{k=k_1}^{k_2}\frac{(-1)^k}{\sin\frac{1}{2}(x-\beta_k)}.$$

Consider

$$\Sigma = \sum_{k=k_1}^{k_2}\frac{(-1)^k}{\sin\frac{1}{2}(x-\beta_k)} = (-1)^{k_1}\sum\left\{\frac{1}{\sin\frac{1}{2}(x-\beta_p)} - \frac{1}{\sin\frac{1}{2}(x-\beta_{p+1})}\right\},$$

where p stands for $k_1 + 2\nu$ and summation extends over all ν in the range $0 \leqq \nu \leqq \frac{1}{2}(k_2 - k_1 - 1)$. Clearly

$$|\Sigma| \leqq \frac{2\pi}{2n+1} \sum \frac{C}{|x-\beta_p||x-\beta_{p+1}|}$$

provided $|x - \beta_p|$ is sufficiently small.

Now, when $|x - t| \geqq 3\Delta_n$ all the β's are on the same side of x and there are by hypothesis an even number of them. Using $|x - t| \geqq 3\Delta_n$ gives

$$|\Sigma| \leqq \frac{C}{2n+1} \sum_{\nu=0}^{\frac{1}{2}(k_2-k_1-1)} \frac{1}{(x-\beta_p)^2} \leqq C \int_{\beta_{k_1}}^{\beta_{k_2}} \frac{du}{(x-u)^2} \leqq \frac{C\Delta_n}{(t-x)^2},$$

and the inequality (7) follows at once.

The third inequality is proved by practically the same argument. This time, however, we must treat the terms in the range $-q/n \leqq t - x \leqq q'/n$ separately. Clearly, by taking q and q' as either 3 or 4 we can arrange so that the ranges $(x + q'/n, x + \Delta_n)$, $(x - \Delta_n, x - q/n)$ contain an even number of β's. For these ranges the above argument applies. As for the range $x - q/n, x + q'/n$,

$$\frac{1}{2n+1} \sum \left| \frac{\sin (n + \frac{1}{2})(x - \beta_k)}{\sin \frac{1}{2}(x - \beta_k)} \right| \leqq C$$

the number of terms in the sum being at most 4.

The proof of the theorem can be completed by familiar arguments. Writing

$$Q_n(u) = |K_n(x, x + u)| + |K_n(x, x - u)|,$$

$$\psi(u) = |f(x + u) - f(x)| + |f(x - u) - f(x)|,$$

we have

$$|I_n(f, x) - f(x)| \leqq \frac{1}{2\pi} \int_0^\pi \psi(u) Q_n(u)\, du$$

$$= \frac{1}{2\pi} \left\{ \int_0^{3/n} + \int_{3/n}^{3\Delta_n} + \int_{3\Delta_n}^{\pi} \right\} = I_1 + I_2 + I_3.$$

Using (6), (7) and (8), we have, at all points of the Lebesgue set,

$$I_1 \leqq Cn \int_0^{3/n} \psi(u)\, du = o(1),$$

$$I_2 + I_3 \leqq \frac{C}{\Delta_n} \int_0^{3\Delta_n} \psi(u)\, du + \frac{C}{n} \int_{3/n}^{\pi} \frac{\psi(u)}{u^2}\, du = \frac{C}{n} \int_{3/n}^{\pi} \frac{\Psi(u)}{u^3}\, du + o(1) = o(1),$$

where $\Psi(u)$ is as usual, the integral of $\psi(u)$. The theorem is proved.

ST. JOHN'S COLLEGE, CAMBRIDGE, ENGLAND.

REMARK ON A RECENT PAPER OF O. ORE

By Guido Zappa

O. Ore, in a recent paper,[1] studies, among other things, the groups he describes as "groups with conformal chains". After examining the principal properties of these groups, he concludes his paper with the following theorem:

Let G be a group such that every subgroup and every quotient group of G have subgroups of every possible order. Then G is a group with conformal chains, and conversely.

Ore notes moreover that it would be of interest to know whether the condition on the quotient groups is necessary. The following theorem proves that the condition on the quotient groups is not necessary, because it is a consequence of the condition on the subgroups.

Theorem. *Let G be a group such that every (proper or improper) subgroup of G has subgroups of every possible order. Let N be a normal subgroup of G. Then every (proper or improper) subgroup of $\bar{G} = G/N$ has subgroups of every possible order.*

Proof. Let g be the order of G. Each group whose order is a prime, or a product of two equal or different primes, has subgroups of every possible order; hence the theorem is true for these groups. We shall prove the theorem by induction with respect to the number of prime factors of g. Suppose the theorem true for each group whose order divides g. The groups G and $\bar{G}$ are multiply isomorphic. Let $\bar{M}$ be a proper subgroup of $\bar{G}$; and let M be the (proper) subgroup of G whose operations correspond to operations of $\bar{M}$ in multiple isomorphism. We have $\bar{M} = M/N$. Since M is in G, M and every subgroup of M have subgroups of every possible order. The conditions of the theorem are then verified for M. Since the order of M divides g, the theorem is true for M, and consequently $\bar{M}$ (and with $\bar{M}$ every proper subgroup of G) has subgroups of every possible order.

It remains to demonstrate that $\bar{G}$ also has subgroups of every possible order. Let n be the order of N. The order of $\bar{G}$ is then $\bar{g} = g/n$. Let $\bar{d}$ be a divisor of $\bar{g}$. We shall prove that $\bar{G}$ has at least a subgroup of order $\bar{d}$. Let p be a prime dividing $\bar{g}/\bar{d}$. Then $\bar{d}$ divides $\bar{g}/p$. If $\bar{G}$ has a subgroup $\bar{S}$ whose order is $\bar{g}/p$, $\bar{S}$ has subgroups of every possible order and hence of order $\bar{d}$; and $\bar{S}$ being in $\bar{G}$, also $\bar{G}$ has subgroups of order $\bar{d}$. It will then be sufficient to demonstrate that if p is a prime dividing $\bar{g}$, $\bar{G}$ has at least a subgroup of order $\bar{g}/p$.

Received November 22, 1939.

[1] Oystein Ore, *Contributions to the theory of groups of finite order*, this Journal, vol. 5 (1939), pp. 431–460.

Since $g = n\bar{g}$, p divides g; hence G has at least a subgroup S of order g/p. If S contains N, $\bar{G}$ has the subgroup $\bar{S} = S/N$, whose order is $\bar{g}/p$, and the theorem is demonstrated. If, on the contrary, S does not contain N, it is easy to prove that $G/N = \bar{G}$ is simply isomorphic to S/I, I being the intersection of N and S. On the other hand, S and every subgroup of S have subgroups of every possible order; and, since the order of S divides g, S/I and $\bar{G}$, simply isomorphic to S/I, have subgroups of every possible order. Hence $\bar{G}$ has a subgroup at least of order p, and the theorem is completely demonstrated.

Consequently, Ore's theorem can be restated as follows:

Let G be a group such that G, and every subgroup of G, has subgroups of every possible order. Then G is a group with conformal chains, and conversely.

Rome, Italy.

GENERALIZED CURVES

By E. J. McShane

The present paper is the first of a sequence of three, the principal object of the sequence being the establishment of existence theorems for Bolza problems in the calculus of variations. These existence theorems will be of sufficient generality to apply to problems with fixed end points and with variable end points, and also to include isoperimetric problems.

The scheme of proof here adopted might be called the method of the auxiliary problem. In very broad outline, it is as follows. The problem is to minimize a functional $f(x)$ on a class S of elements x. To show the existence of a solution, the range of the functional f is extended to a larger class S^* of elements possessing more desirable properties than S; specifically, a kind of local compactness. By use of these properties of S^* it is shown that there exists an element x_0 of S^* which minimizes $f(x)$ on S^*. Next we find the necessary conditions which are satisfied by x_0 as a consequence of minimizing $f(x)$. Under suitable hypotheses on f these conditions will imply that x_0 is a member, not merely of the extended class S^*, but of the original class S. Thus x_0 is the minimizing element which is the solution of the original problem.

In this note and its two successors we shall study single-integral problems by introducing an auxiliary problem. The original class S of curves will be enlarged to the class S^* of generalized curves, invented by L. C. Young;[1,2,3] and each of the three papers of the sequence will develop one of the three steps outlined in the general proof-pattern. In this first paper we develop the theory of generalized curves, and show that they possess a compactness property which leads us readily to an existence theorem in the class of generalized curves. In the second paper we shall develop the theory of the calculus of variations as extended to generalized curves, obtaining (for Bolza problems) analogues of the multiplier rule, the Weierstrass and Clebsch conditions, and the Dresden corner condition. In the third paper we set forth additional hypotheses on the integrands involved which guarantee that the minimizing generalized curve is actually a curve in the ordinary sense. In all three papers we shall consider only problems in parametric form.

I wish here to express my thanks to Professor L. M. Graves, who read the original manuscripts and suggested a number of improvements; for instance, the

Received November 28, 1939.

[1] L. C. Young, *On approximation by polygons in the calculus of variations*, Proceedings of the Royal Society, (A), vol. 141(1933), pp. 325–341.

[2] L. C. Young, *Generalized curves and the existence of an attained absolute minimum in the calculus of variations*, Comptes Rendus de la Société des Sciences et des Lettres de Varsovie, Classe III, vol. 30(1937), pp. 212–234.

[3] L. C. Young, *Necessary conditions in the calculus of variations*, Acta Math., vol. 69(1938), pp. 239–258.

replacement of the previous definition of generalized curve by an equivalent formulation (§2) more appropriate for use here.

1. **Heuristic discussion.** If we are trying to establish an existence theorem for, say, an isoperimetric problem, we naturally must consider the class of rectifiable curves (or some subclass) and the integrals, of the usual type of the calculus of variations, taken along such curves. There are obvious advantages in topologizing the space of rectifiable curves in such a way that the integrals along them will be continuous. This can be done in many ways; but, to reduce the troubles of finding a convergent minimizing sequence, we shall choose the weakest topology in which all integrals are continuous. In other words, we shall say that a sequence $\{C_n\}$ has C_0 for limit if and only if $\mathcal{F}(C_n)$ tends to $\mathcal{F}(C_0)$ for every integral $\mathcal{F}(C)$. The space of rectifiable curves is thus given a topology; in fact, it is even a metric space. However, it is not complete. Less precisely expressed, it would clearly be an advantage to us if for every sequence $\{C_n\}$ with the property that $\{\mathcal{F}(C_n)\}$ converges for every integral $\mathcal{F}(C)$ of the calculus of variations, there would exist a curve C_0 such that $C_n \to C_0$. This is not the case; so, as usual, we invent new entities to fill the gaps. These new entities are the generalized curves of L. C. Young.

Consider, for example, the sequence of curves

$$(1.1) \qquad x = x_n(t) \equiv t, \qquad y = y_n(t) \qquad (0 \leqq t \leqq 1),$$

in which $y_n(t)$ vanishes at $0, 2/2n, 4/2n, \cdots, 1$, is equal to $1/2n$ at $1/2n, 3/2n, \cdots, (2n-1)/2n$, and is linear on each interval $[p/2n, (p+1)/2n]$, so that y_n' is alternately $+1$ and -1. If $F(x, y, x', y')$ is any integrand, we easily compute

$$(1.2) \qquad \lim_{n\to\infty} \int_0^1 F(x_n, y_n, x_n', y_n')\, dt = \int_0^1 \tfrac{1}{2}\{F(t, 0, 1, 1) + F(t, 0, 1, -1)\}\, dt.$$

If the curves (1.1) are to approach a limit, the integral along this limit must therefore equal the right member of (1.2). The formation of this integral may be thus described. In the integrand function $F(x, y, x', y')$ we first substitute for x, y the limits

$$x_0(t) = t, \qquad y_0(t) \equiv 0$$

of the functions (1.1); this gives a function $F(x_0(t), y_0(t), x', y')$ of t, x', and y'. Next, for each t we form a certain average of values of $F(x_0(t), y_0(t), x', y')$; namely, the arithmetic mean of the values at $x' = 1$, $y' = 1$ and at $x' = 1$, $y' = -1$. The right member of (1.2) is then the integral of this average.

This would suggest that the notion of generalized (plane) curve C_0^* should contain two essential elements; first, a pair of functions $x_0(t)$, $y_0(t)$; secondly, a description of an averaging process to be carried out on the values of $F(x_0(t), y_0(t), x', y')$ in order to compute the number which can be called the integral of F along C_0^*. These two elements are not independent of each other. In the

specific example under discussion, if we choose $F(x, y, x', y') = x'$, we find that the average of the values of F is 1, which is x_0'; likewise the average of the values of y' is 0, which is y_0'. A moment's reflection will show that this is not fortuitous; such a connection must exist in order for the integrals of x' and y' to be continuous.

2. Definition of generalized curve. Before setting forth our definition of generalized curve it is desirable to introduce some notation.

The letters y, r (with or without suffixes) will stand for vectors in ν-dimensional space: $y = (y^1, \cdots, y^\nu)$, $r = (r^1, \cdots, r^\nu)$. The repetition of a *Greek-letter* affix other than ν requires summation over all values of that affix; thus

$$y_n^\alpha r_n^\alpha = y_n^1 r_n^1 + \cdots + y_n^\nu r_n^\nu .$$

The length of a vector is denoted by enclosing it between vertical bars; thus

$$| r | = [r^\alpha r^\alpha]^{\frac{1}{2}}.$$

Let Q be the aggregate of functions $\Phi(r)$ defined and continuous for all r. Let $\mathfrak{M}[\Phi]$ be a functional which assigns a real number to each element Φ of the class Q. The functional $\mathfrak{M}[\Phi]$ is

(2.1) *distributive*, if for all real numbers a_1, a_2 and all pairs of elements Φ_1, Φ_2 of Q the equation $\mathfrak{M}[a_1\Phi_1 + a_2\Phi_2] = a_1\mathfrak{M}[\Phi_1] + a_2\mathfrak{M}[\Phi_2]$ holds;

(2.2) *non-negative*, if $\mathfrak{M}[\Phi] \geqq 0$ for all non-negative functions $\Phi(r)$ in Q;

(2.3) *determined by the values of Φ on a set R*, if for every pair Φ_1, Φ_2 of functions of Q which coincide on R the equation $\mathfrak{M}[\Phi_1] = \mathfrak{M}[\Phi_2]$ holds.

The following remarks are evident.

Lemma 2.1. *If $\mathfrak{M}[\Phi]$ is distributive, it is determined by the values of Φ on R if and only if $\mathfrak{M}[\Phi] = 0$ for every continuous function which vanishes on R.*

Lemma 2.2. *If $\mathfrak{M}[\Phi]$ is distributive and non-negative and determined by the values of Φ on a set R, then*

(2.4) $\mathfrak{M}[\Phi_1] \leqq \mathfrak{M}[\Phi_2]$ *provided that* $\Phi_1(r) \leqq \Phi_2(r)$ *on* R;

(2.5) $| \mathfrak{M}[\Phi] | \leqq \mathfrak{M}[| \Phi |]$;

(2.6) $| \mathfrak{M}[\Phi] | \leqq \mathfrak{M}[1] \{\text{l.u.b. } | \Phi(r) | \text{ on } R\}$;

(2.7) *if Φ_0, Φ_1, $\cdots$ are functions of Q, and Φ_n converges to Φ uniformly on R, then*

$$\lim_{n\to\infty} \mathfrak{M}[\Phi_n] = \mathfrak{M}[\Phi].$$

We can now write our definition of generalized curve.[4]

(2.8) *A generalized curve is a system consisting of a set of functions $y^i(t)$ $(a \leqq t \leqq b;$ $i = 1, \cdots, \nu)$, a subset M of the interval $[a, b]$ with measure $b - a$, and a functional $\mathfrak{M}[t; \Phi]$ defined for all t in M and all continuous $\Phi(r)$, and possessing the following properties.*

[4] L. C. Young, loc. cit. (footnote 2), p. 231.

(2.8a) *The functions $y^i(t)$ are absolutely continuous on $[a, b]$ and possess finite derivatives on M.*
(2.8b) *For each t in M, the functional $\mathfrak{M}[t; \Phi]$ is distributive and non-negative, and is determined by the values of $\Phi(r)$ on a bounded set of values of r.*
(2.8c) *For each t in M,*

$$\mathfrak{M}[t; r^i] = y^{i\prime}(t) \qquad (i = 1, \cdots, \nu).$$

(2.8d) *For each continuous function $\Phi(r)$, the functional $\mathfrak{M}[t; \Phi]$ is measurable on M.*

In (2.8c) the notation may be slightly confusing. The functional $\mathfrak{M}[t; \Phi]$ does not depend on a number r^i; on the contrary, by $\mathfrak{M}[t; r^i]$ we mean the number $\mathfrak{M}[t; \Phi]$ which is determined when for Φ we choose the particular continuous function defined by the equation $\Phi(r^1, \cdots, r^n) = r^i$. Occasionally we shall meet this situation again; we shall write $\mathfrak{M}[t; \Phi(r)]$, but shall always mean the value of $\mathfrak{M}[t; \Phi]$ at some particular function $\Phi(r)$.

The generalized curve defined in (2.8) will be denoted by

$$C^*: [y(t), \mathfrak{M}[t; \Phi], M]. \tag{2.9}$$

In this notation the interval $[a, b]$ can be thought of as the closure $\bar{M}$ of the set M. The ordinary curve $y = y(t)$ $(a \leqq t \leqq b)$ will be called the *track* of C^*.

3. Integrals along generalized curves. We shall usually be interested in the generalized curves whose tracks lie in some particular closed set E, and in integrals whose integrands $F(y, r)$ are defined and continuous for all y in E and all r. However, we can avoid verbosity by always supposing that the range of F has been extended so that F is defined and continuous for all y and all r. If F is positively homogeneous of degree 1 in r, so that

$$F(y, kr) = kF(y, r), \qquad k \geqq 0, \tag{3.1}$$

for all y in E, we can first find a continuous function $\Phi(y, r)$ which coincides with $F(y, r)$ when y is in E, and we can then define

$$\Psi(y, r) = \Phi(y, r/|r|)\cdot|r| \quad \text{if} \quad |r| \neq 0,$$

$$\Psi(y, 0) = 0.$$

The function Ψ is easily seen to be defined and continuous for all (y, r) and to be positively homogeneous of degree 1 in r. Also, if y is in E, the functions $\Psi(y, r)$ and $F(y, r)$ are equal. Thus
(3.2) *there is no loss of generality in supposing that all integrand functions are defined and continuous for all (y, r).*
For brevity, we shall say that
(3.3) *$F(y, r)$ is a "parametric integrand" if it is defined and continuous for all y and all r, and is positively homogeneous of degree 1 in r.*

Given a parametric integrand $F(y, r)$ and a generalized curve (2.9), we define

$$\mathcal{F}(C^*) = \int_M \mathfrak{M}[t; F(y(t), r)]\, dt, \tag{3.4}$$

provided that the integral exists. It is usually more convenient to write

$$\mathcal{F}(C^*) = \int_a^b \mathfrak{M}[t; F(y(t), r)]\, dt, \tag{3.5}$$

the integrand being given the value zero on the complement of M.

By the definition (2.8), the function $\mathfrak{M}[t; \mid r \mid]$ is non-negative and measurable on M. If it is summable, we say that C^* is *rectifiable*, or has *finite length*, and we define its length to be

$$\mathfrak{L}(C^*) = \int_a^b \mathfrak{M}[t; \mid r \mid]\, dt. \tag{3.6}$$

If C^* has finite length, so has its track; and the two lengths satisfy the inequality

$$\int_a^b \mid y'(t) \mid dt \leqq \int_a^b \mathfrak{M}[t; \mid r \mid]\, dt. \tag{3.7}$$

For by the Cauchy-Schwarz inequality we have $r^i y^{i\prime} \leqq \mid r \mid \cdot \mid y' \mid$. Hence, by (2.3b, c), for t in M we have

$$y^{i\prime} y^{i\prime} \leqq \mid y' \mid \mathfrak{M}[t; \mid r \mid],$$

so that $\mid y' \mid \leqq \mathfrak{M}[t; \mid r \mid]$. Integration yields (3.7).

The next lemma expresses an important property of rectifiable generalized curves.

LEMMA 3.1. *If C^* is a rectifiable generalized curve, and $\Phi(t, r)$ is defined and continuous for all t in $[a, b]$ and all r, the function $\mathfrak{M}[t; \Phi(t, r)]$ is measurable on M.*

Let $[a, b]$ be subdivided into n equal subintervals $\Delta_{n,1}, \cdots, \Delta_{n,n}$ by points $t_{n,0} = a, t_{n,1}, \cdots, t_{n,n} = b$. Define

$$\Phi_n(t, r) = \sum_{i=1}^{n} \Phi(t_{n,i}, r)\chi_{n,i}(t),$$

where $\chi_{n,i}(t)$ is the characteristic function of the interval $\Delta_{n,i}$. Then

$$\mathfrak{M}[t; \Phi_n(t, r)] = \sum_{i=1}^{n} \mathfrak{M}[t; \Phi(t_{n,i}, r)]\, \chi_{n,i}(t)$$

is measurable on M, by (2.8d). For each t_0 in M the value of $\mathfrak{M}[t_0; \Phi]$ is determined by the values of Φ on some sphere $\mid r \mid \leqq k$, and on this sphere the functions $\Phi_n(t_0, r)$ tend uniformly to $\Phi(t_0, r)$. By Lemma 2.2, the measurable functions $\mathfrak{M}[t; \Phi_n(t, r)]$ tend on M to $\mathfrak{M}[t; \Phi(t, r)]$, which is therefore measurable on M.

LEMMA 3.2. *If C^* is a rectifiable generalized curve and $F(y, r)$ is a parametric integrand, the integral $\mathcal{F}(C^*)$ of equation (3.4) is defined.*

Let K be an upper bound for the absolute value of $F(y(t), r)$ on the bounded closed set $[a \leqq t \leqq b, |r| = 1]$. By homogeneity we have

$$|F(y(t), r)| \leqq K|r| \qquad (a \leqq t \leqq b).$$

Hence by Lemma 2.2

$$(3.8) \qquad |\mathfrak{M}[t; F(y(t), r)]| \leqq K\mathfrak{M}[t; |r|]$$

for all t in M. Since by Lemma 3.1 the function $\mathfrak{M}[t; F(y(t), r)]$ is measurable on M, inequality (3.8) proves it summable.

4. **A topology in the space of generalized curves.** In the introduction we have mentioned the topology which we shall use.

(4.1) *The rectifiable generalized curve C_0^* is the limit of the sequence $\{C_n^*\}$ of rectifiable generalized curves if*

(a) *the Fréchet distance between the track of C_0^* and that of C_n^* tends to zero as $n \to \infty$, and*

(b) *for every parametric integrand $F(y, r)$ it is true that*

$$\lim_{n\to\infty} \mathcal{F}(C_n^*) = \mathcal{F}(C_0^*).$$

Correspondingly, we shall adopt the following definition of identity of generalized curves.

(4.2) *The generalized curves C_0^* and C_1^* are identical if*

(a) *their tracks are identical (i.e., have Fréchet distance zero), and*

(b) $\mathcal{F}(C_0^*) = \mathcal{F}(C_1^*)$ *for every parametric integrand $F(y, r)$.*

The conditions (4.1a) and (4.2a) have been added for convenience in our proofs. It could, however, be shown that they can be greatly weakened, and if C_0^* is non-degenerate, they can be entirely omitted without changing the theorems later to be proved.

With these definitions, the space of generalized curves is an $\mathfrak{L}^*$-space, in the terminology of Kuratowski. It is in fact more than that; it is metrizable. However, this fact is of no immediate use to us, so we relegate its proof to the appendix (§12).

The convention (4.2) of identity permits us to prove the following lemma.

Lemma 4.1. *If C^* is a rectifiable generalized curve* (2.9), *and k is a positive number, we may suppose without loss of generality that for all t in M the functional $\mathfrak{M}[t; \Phi]$ depends only on the values of $\Phi(r)$ on the sphere $|r| = k$.*

If $\Phi(y, r)$ is any function continuous in (y, r), we define $\Phi_k(y, r)$ to be the function which is positively homogeneous of degree 1 in r and coincides with $\Phi(y, r)$ when $|r| = k$:

$$(4.3) \qquad \begin{aligned} \Phi_k(y, r) &= k^{-1}|r|\,\Phi(y, kr/|r|), \qquad |r| > 0. \\ \Phi_k(y, 0) &= 0. \end{aligned}$$

This is easily seen to be continuous. Moreover, if $\Phi(y, r)$ is already positively homogeneous of degree 1 in r, the identity

$$\Phi_h(y, r) \equiv \Phi(y, r) \tag{4.4}$$

holds.

We now define C_0^* by the formula $[y(t), \mathfrak{M}_0[t; \Phi], M]$, where

$$\mathfrak{M}_0[t; \Phi] = \mathfrak{M}[t; \Phi_h]. \tag{4.5}$$

Since Φ_h is determined by the values of Φ on the sphere $| r | = k$, so is $\mathfrak{M}_0[t; \Phi]$. The system C_0^* evidently satisfies (2.8a, b); by (4.5) it satisfies (2.8d), since Φ_h is continuous; and by (4.4) it satisfies (2.8c). Thus C_0^* is a generalized curve. The curves C^* and C_0^* have the same track, and by (4.4) we have $\mathcal{F}(C_0^*) = \mathcal{F}(C^*)$ for every parametric integrand; so C_0^* is identical with C^*. This completes the proof.

Let us say that

(4.6) *a generalized curve* C^*: $[y(t), \mathfrak{M}[t; \Phi], M]$ *is an isomorph of an ordinary curve if for almost all* t *in* M *the functional* $\mathfrak{M}[t; \Phi]$ *depends only on the value of* Φ *at a single spot* r_t.

We can then prove

LEMMA 4.2. *If the rectifiable ordinary curves C and the isomorphs C^* of ordinary curves are set into correspondence by letting each C^* correspond to its own track, the correspondence is one-to-one. Moreover, for every parametric integrand $F(y, r)$ we have*

$$\int_a^b F(y(t), \dot{y}(t))\, dt = \int_a^b \mathfrak{M}[t; F(y(t), r)]\, dt. \tag{4.7}$$

If $\mathfrak{M}[t; \Phi]$ depends only on the value of $\Phi(r)$ at r_t, by (2.8b) it has the form

$$\mathfrak{M}[t; \Phi] = a\Phi(r_t), \qquad a \geqq 0. \tag{4.8}$$

By (2.8c) this implies

$$a r_t^i = y^{i\prime}(t) \qquad (i = 1, \cdots, \nu). \tag{4.9}$$

If $F(y, r)$ is a parametric integrand, equations (4.8) and (4.9) yield

$$\begin{aligned} \mathfrak{M}[t; F(y(t), r)] &= aF(y(t), r_t) \\ &= F(y(t), y'(t)). \end{aligned} \tag{4.10}$$

This holds for almost all t in $[a, b]$, so (4.7) follows by integration.

If two isomorphs C_1^*, C_2^* of generalized curves both correspond to $C: y = y(t)$ $(a \leqq t \leqq b)$, the curve C is the track of both ordinary curves, and by (4.7)

$$\mathcal{F}(C_1^*) = \mathcal{F}(C_2^*)$$

for all integrands $F(y, r)$. Thus by (4.2) the curves are identical. This completes the proof of the lemma.

Thus we see that our generalized curves form an extension of the class of ordinary curves in the same way that, for example, the real numbers form an extension of the class of rationals.

5. **Elementary change of parameter.** Let C^*: $[y_1(\tau), \mathfrak{M}_1[\tau; \Phi], M_1]$ be a generalized curve, wherein $\bar{M}_1$ is the interval $[a_1, b_1]$, and let $\tau = \tau(t)$ be a function defined and absolutely continuous on the interval $a \leqq t \leqq b$ and satisfying the conditions

$$\tau'(t) > 0 \quad \text{for almost all } t, \tag{5.1}$$

$$\tau(a) = a_1, \quad \tau(b) = b_1. \tag{5.2}$$

We wish to define the concept of a change of parameter on C^* from τ to t; such a change, under the above hypotheses on $\tau(t)$, will be called an *elementary change of parameter*.

The new equation of the track will be

$$y = y(t) = y_1(\tau(t)) \qquad (a \leqq t \leqq b); \tag{5.3}$$

the functions $y(t)$ are then absolutely continuous. Because of (5.1) the function $\tau(t)$ has an absolutely continuous inverse,[5] so for almost all t in $[a, b]$ the point $\tau(t)$ is in M_1. Let M be the subset of $[a, b]$ for which $\tau'(t)$ is defined and finite and $\tau(t)$ is in M_1. This set has measure $b - a$, and on it

$$y'(t) = y_1'(\tau(t))\cdot\tau'(t). \tag{5.4}$$

Thus (2.8a) is verified.

For t in M we define

$$\mathfrak{M}[t; \Phi] = \mathfrak{M}_1[\tau(t); \Phi]\cdot\tau'(t). \tag{5.5}$$

If we set $\Phi(r) = r^j$, equations (5.5), (5.4) and (2.8c) give $\mathfrak{M}[t; r^j] = y'(t)$, so the new system satisfies (2.8c). Condition (2.8b) is evident. For (2.8d) we can either use the facts that $\tau(t)$ has an absolutely continuous inverse and that τ' is measurable, which with (5.5) give the measurability of $\mathfrak{M}[t; \Phi]$; or we can use the theorem on change of variables,

$$\begin{aligned} \int_{a_1}^{b_1} \mathfrak{M}_1[\tau; \Phi(\tau, r)]\, d\tau &= \int_a^b \mathfrak{M}_1[\tau(t); \Phi(\tau(t), r]\, \tau'(t)\, dt \\ &= \int_a^b \mathfrak{M}[\tau(t); \Phi(\tau(t), r)]\, dt. \end{aligned} \tag{5.6}$$

This gives (2.8d) if Φ is independent of τ, and it also informs us that every integral $\mathcal{F}(C)$ is invariant under elementary change of parameter. This invariance, together with equation (5.3), which states that the tracks $y = y_1(\tau)$ $(a_1 \leqq \tau \leqq b_1)$ and $y = y(t)$ $(a \leqq t \leqq b)$ are identical, assures us that the condi-

[5] C. Carathéodory, *Vorlesungen über reele Funktionen*, p. 584, Satz 3.

tion (4.2) for identity is satisfied, so that an elementary change of parameter gives us back the same curve in a new representation.

The following lemma will be useful in proving compactness.

LEMMA 5.1. *If*

$$C^*: [y_1(\tau), \mathfrak{M}_1[\tau; \Phi], M_1]$$

is a rectifiable generalized curve, for every positive ϵ *the curve* C^* *has a representation*

$$C^*: [y(t), \mathfrak{M}[t; \Phi], M]$$

such that $\bar{M}$ *is the interval* [0, 1] *and*

$$\mathfrak{M}[t; |r|] \leqq \mathfrak{L}(C^*) + \epsilon \tag{5.7}$$

for all t *in* M.

Denote $\bar{M}_1$ by $[a, b]$, and define

$$t(\tau) = \{\mathfrak{L}(C^*) + \epsilon\}^{-1} \int_a^\tau \{\mathfrak{M}_1[\tau; |r|] + \epsilon(b-a)^{-1}\}\, d\tau.$$

Then the image of the interval $a \leqq \tau \leqq b$ is $0 \leqq t \leqq 1$. The elementary change of parameter from τ to t gives, by (5.5),

$$\mathfrak{M}[t; \Phi] = \frac{\mathfrak{M}_1[\tau(t); \Phi]\{\mathfrak{L}(C^*) + \epsilon\}}{\mathfrak{M}_1[\tau(t); |r|] + \epsilon(b-a)^{-1}},$$

valid for all t in the set M of measure 1. For $\Phi(r) = |r|$ this yields (5.7).

6. A compactness theorem.

The following theorem is an obvious consequence of (4.1b).

THEOREM 6.1. *If* $F(y, r)$ *is a parametric integrand, then the integral* $\mathfrak{F}(C^*)$, *regarded as a functional on the space of rectifiable generalized curves, is continuous.*

L. C. Young[6] has proved a theorem of considerable interest in itself, which we shall merely state without proof, since we do not use it. In our terminology, it is as follows.

THEOREM 6.2. *The set of isomorphs of ordinary curves is everywhere dense in the space of rectifiable generalized curves.*

Our next theorem[7] states a compactness property essential in the later proofs of existence theorems.

THEOREM 6.3. *Let* E_0 *be a bounded closed set in* y*-space, and let* N *be a number. The set of generalized curves* C^* *whose tracks lie in* E_0 *and whose lengths do not exceed* N *is a compact set.*

[6] Loc. cit. (footnote 2), pp. 225, 233.

[7] This is closely related to a theorem of L. C. Young (see p. 224 of the reference in footnote 2).

Let $C_1^*, C_2^*, \cdots$ be a sequence of generalized curves satisfying the hypothesis. We must show that there exists a subsequence which converges to a limit curve C_0^*; this limit curve will evidently have its track in E_0, since E_0 is closed, and will have length at most N, since the length-integral is continuous by Theorem 6.1.

For each n we introduce on C_n^* the parameter of Lemma 5.1, with $\epsilon = n^{-1}$. Then C_n^* has the representation

$$C_n^* : [y_n(t), \mathfrak{M}_n[t; \Phi], M_n],$$

where $\bar{M}_n$ is the interval [0, 1] and

$$\mathfrak{M}_n[t; | r |] \leqq \mathfrak{L}(C_n^*) + n^{-1} \leqq N + 1 \tag{6.1}$$

for all t in M_n. We also suppose, as by Lemma 4.1 we may, that

(6.2) *for all t in M_n the value of $\mathfrak{M}_n[t; \Phi]$ is determined by the values of $\Phi(r)$ on the set $| r | = 1$.*

First we choose a subsequence of $\{C_n\}$, retaining the notation of the original sequence, such that

(6.3) *$y_n(0)$ converges to some limit as n tends to infinity.*

This is possible because of the boundedness of E_0. Let $\Phi(y, r)$ be an arbitrary function continuous on (y, r)-space. By Lemma 3.1, the functions $\mathfrak{M}_n[t; \Phi(y_n(t), r)]$ are measurable on the respective sets M_n. If H is an upper bound for the absolute value of Φ on the bounded closed set $[y \text{ in } E_0, | r | = 1]$, by Lemma 2.2 and (6.1) we have

$$| \mathfrak{M}_n[t; \Phi(y_n(t), r)] | \leqq H\mathfrak{M}_n[t; | r |] \leqq H(N + 1).$$

Hence the integrals

$$\varphi_n(t) = \int_0^t \mathfrak{M}_n[t; \Phi(y_n(t), r)]\, dt \tag{6.4}$$

are defined and satisfy the inequality

$$| \varphi_n(t_2) - \varphi_n(t_1) | = \left| \int_{t_1}^{t_2} \mathfrak{M}_n[t; \Phi(y_n(t), r)]\, dt \right| \leqq H(N + 1) | t_2 - t_1 | \tag{6.5}$$

if t_1 and t_2 are in [0, 1]. So the functions $\varphi_n(t)$ all satisfy the same Lipschitz condition; and since $\varphi_n(0) = 0$ they are uniformly bounded. By Ascoli's theorem, there is a subsequence (we continue to use the notation of the original sequence) which converges uniformly to a limit function $\varphi_0(t)$:

$$\lim_{n \to \infty} \varphi_n(t) = \varphi_0(t), \qquad \text{uniformly on } [0, 1]. \tag{6.6}$$

It is possible to find a denumerable collection $[\Phi_m(y, r)]$ of continuous functions such that for every continuous function $\Phi(y, r)$ and every positive ϵ there is a Φ_m of the collection which differs from Φ by less than ϵ on the set $[y \text{ in } E_0, | r | = 1]$. We prefer to write this in the equivalent form

$$| \Phi(y, r) - \Phi_m(y, r) | < \epsilon | r | \qquad (y \text{ in } E_0, | r | = 1). \tag{6.7}$$

From the preceding paragraph we see that, by use of the diagonal process, we can select a subsequence of $\{C_n^*\}$ (we continue to use the notation $\{C_n^*\}$ for the subsequence) such that for each m the sequence of functions

$$\varphi_{n,m}(t) = \int_0^t \mathfrak{M}_n[t; \Phi_m(y_n(t), r)]\, dt \tag{6.8}$$

converges uniformly as $n \to \infty$ to a Lipschitzian limit function $\varphi_{0,m}(t)$.

For each m the function $\varphi_{0,m}(t)$ has a derivative except on a set of measure zero. The sum of these exceptional sets is still of measure zero; discarding it, we are left with a set M of measure 1 on which all the functions $\varphi_{0,m}(t)$ are differentiable.

Let $\Phi(y, r)$ be a continuous function. Given a positive ϵ, we choose an m for which inequality (6.7) holds. From the uniform convergence of the $\varphi_{n,m}(t)$, there is an n_0 such that if p and q both exceed n_0 the inequality

$$| \varphi_{p,m}(t) - \varphi_{q,m}(t) | < \epsilon \tag{6.9}$$

holds for all t in $[0, 1]$. From (6.4) and (6.7), with Lemma 2.2, the inequality

$$\begin{aligned} | \varphi_{n,m}(t) - \varphi_n(t) | &\leqq \int_0^t | \mathfrak{M}_n[t; \Phi_m(y_n(t), r)] - \mathfrak{M}_n[t; \Phi(y_n(t), r)] |\, dt \\ &\leqq \int_0^1 \mathfrak{M}_n[t; | \Phi_m(y_n, r) - \Phi(y_n(t), r) |]\, dt \\ &\leqq \epsilon \int_0^1 \mathfrak{M}_n[t; | r |]\, dt \leqq \epsilon N \end{aligned} \tag{6.10}$$

holds for all n. Combining (6.9) and (6.10) yields

$$| \varphi_p(t) - \varphi_q(t) | < (2N + 1)\epsilon, \tag{6.11}$$

whenever p and q both exceed n_0. This implies the uniform convergence of $\varphi_n(t)$ to a limit function $\varphi_0(t)$. Thus for our chosen subsequence we have the simultaneous uniform convergence of all indefinite integrals of the form (6.4) to limit functions.

Again, let $\Phi(y, r)$ be an arbitrary continuous function. For an arbitrary positive number ϵ we select an m such that inequality (6.7) holds. Then the functions

$$\Phi(y, r) - \Phi_m(y, r) - \epsilon | r |, \qquad \Phi(y, r) - \Phi_m(y, r) + \epsilon | r | \tag{6.12}$$

are respectively negative and positive for y in E_0 and $| r | = 1$. With the help of Lemma 2.2, statements (6.1) and (6.12) imply that

$$\begin{aligned} &\mathfrak{M}_n[t; \Phi(y_n(t), r)] - \mathfrak{M}_n[t; \Phi_m(y_n(t), r)] - \epsilon(N + 1) \\ &\qquad < 0 < \mathfrak{M}_n[t; \Phi(y_n(t), r)] - \mathfrak{M}_n[t; \Phi_m(y_n(t), r)] + \epsilon(N + 1). \end{aligned} \tag{6.13}$$

On integrating and using definitions (6.4) and (6.8), this shows that the functions

$$\varphi_n(t) - \varphi_{n,m}(t) - \epsilon(N + 1)t, \qquad \varphi_n(t) - \varphi_{n,m}(t) + \epsilon(N + 1)t \tag{6.14}$$

are respectively non-increasing and non-decreasing. Passing to the limit, we see that the same is true of the functions

$$(6.15) \qquad \varphi_0(t) - \varphi_{0,m}(t) - \epsilon(N+1)t, \qquad \varphi_0(t) - \varphi_{0,m}(t) + \epsilon(N+1)t.$$

Therefore the first of functions (6.15) has a non-positive upper derivative, and the second has a non-negative lower derivative. In particular, since $\varphi_{0,m}(t)$ has a derivative at t_0, we have

$$(6.16) \quad \bar{D}\varphi_0(t_0) - \varphi_{0,m}'(t_0) - \epsilon(N+1) \leqq 0 \leqq \underline{D}\varphi_0(t_0) - \varphi_{0,m}'(t_0) + \epsilon(N+1).$$

From this inequality we have at once

$$(6.17) \qquad 0 \leqq \bar{D}\varphi_0(t_0) - \underline{D}\varphi_0(t_0) \leqq 2\epsilon(N+1).$$

This holds for all positive ϵ, so the upper and lower derivatives of φ_0 are equal' and $\varphi_0'(t_0)$ exists. That is,

(6.18) *for our chosen subsequence* $\{C_n^*\}$ *and the fixed set* M *of measure* 1, *every function* $\varphi_0(t)$ *defined by* (6.6) *is differentiable at every point of* M.

We are now ready to define our limit curve C_0^*. First, if we apply the foregoing to the function $\Phi(r) = r^j$, we find by (2.8c)

$$\varphi_n(t) = y_n^j(t) - y_n^j(0).$$

So by the preceding proof, together with statement (6.3), the $y_n^j(t)$ converge uniformly to a Lipschitzian limit $y_0^j(t)$, and this limit function is differentiable at each point of M. We shall use the curve $y = y_0(t)$ $(0 \leqq t \leqq 1)$ for the track of C_0^*; condition (2.8a) is then satisfied, by (6.18). The functional $\mathfrak{M}_0[t; \Phi]$ will be defined to be the derivative at t of the function $\varphi_0(t)$ defined by (6.4) and (6.6). That is, for each t in M and each continuous function $\Phi(r)$ we define

$$(6.19) \qquad \mathfrak{M}_0[t; \Phi] = \frac{d}{dt} \lim_{n \to \infty} \int_0^t \mathfrak{M}_n[t; \Phi]\, dt.$$

If we substitute r^j for Φ, we find that (2.8c) is satisfied. There is no difficulty in verifying (2.8b); and (2.8d) is satisfied because $\mathfrak{M}_0[t; \Phi]$ is the derivative of an absolutely continuous function. Hence the system

$$C_0^* : [y_0(t), \mathfrak{M}_0[t; \Phi], M]$$

is a generalized curve.

Let $\Phi(y, r)$ be an arbitrary continuous function, and let t_0 be a point in M. Define

$$\Psi(r) = \Phi(y_0(t_0), r).$$

If ϵ is an arbitrary positive number, there is a neighborhood U of $y_0(t_0)$ on which

$$(6.20) \qquad | \Phi(y, r) - \Psi(r) | < \epsilon | r |, \qquad | r | = 1.$$

If n is large enough and (α, β) is a sufficiently small interval containing t_0, the arcs of track $y = y_n(t)$ $(\alpha \leqq t \leqq \beta)$ lie in U. Hence by (6.20) and (6.1)

$$(6.21) \qquad \left| \int_\alpha^\beta \mathfrak{M}_n[t; \Phi(y_n(t), r)]\, dt - \int_\alpha^\beta \mathfrak{M}_n[t; \Psi(r)]\, dt \right| \leqq \epsilon(N + 1)(\beta - \alpha),$$

or with notation (6.4)

$$| \{\varphi_n(\beta) - \varphi_n(\alpha)\} - \{\psi_n(\beta) - \psi_n(\alpha)\} | \leqq \epsilon(N + 1)(\beta - \alpha).$$

Passing to the limit, we get

$$(6.22) \qquad | \{\varphi_0(\beta) - \varphi_0(\alpha)\} - \{\psi_0(\beta) - \psi_0(\alpha)\} | \leqq \epsilon(N + 1)(\beta - \alpha).$$

If we divide by $\beta - \alpha$ and recall that ϵ is arbitrary, this shows that the quotients

$$\frac{\varphi_0(\beta) - \varphi_0(\alpha)}{\beta - \alpha} \quad \text{and} \quad \frac{\psi_0(\beta) - \psi_0(\alpha)}{\beta - \alpha}$$

have the same limit as α and β tend to t_0. But by definition the limit of the latter of these is $\mathfrak{M}_0[t_0 ; \Phi(y_0(t_0), r)]$. Hence for all t_0 in M we have

$$(6.23) \qquad \varphi_0'(t_0) = \mathfrak{M}_0[t_0 ; \Phi(y_0(t_0), r)],$$

where φ_0 is defined by (6.4) and (6.6). It follows at once that

$$(6.24) \qquad \begin{aligned} \lim_{n\to\infty} \mathfrak{F}(C_n^*) &= \lim_{n\to\infty} \varphi_n(1) = \varphi_0(1) \\ &= \int_0^1 \dot\varphi_0(t)\, dt = \int_0^1 \mathfrak{M}_0[\tau; \Phi(y_0(t), r)]\, dt = \mathfrak{F}(C_0^*). \end{aligned}$$

Since the track-functions $y_n(t)$ converge uniformly to $y_0(t)$, this shows that C_n^* converges to C_0^*, and the theorem is established.

7. **Standard parameter.** Although it is not necessary for our proofs, it is convenient to know that every rectifiable generalized curve has a representation in which the parameter is proportional to length of generalized arc.

LEMMA 7.1. *If the generalized curve*

$$(7.1) \qquad C^*: [y_1(t), \mathfrak{M}_1[t; \Phi], M_1]$$

is rectifiable, it has a representation

$$(7.2) \qquad C^*: [y(t), \mathfrak{M}[t; \Phi], M]$$

with the following properties.

(7.3) *$\bar M$ is the interval* $0 \leqq t \leqq 1$.

(7.4) *For all t in M, the functional $\mathfrak{M}[t; \Phi]$ depends only on the values of Φ on the sphere* $| r | = 1$.

(7.5) *For all t in M,*

$$\mathfrak{M}[t; | r |] = \mathfrak{L}(C^*).$$

(7.6) *For every continuous function* $\Phi(t, r)$ *the integral of* $\mathfrak{M}[t; \Phi(t, r)]$ *is differentiable at each* t *in* M, *and*

$$\frac{d}{dt}\int_0^t \mathfrak{M}[t; \Phi(t, r)]\,dt = \mathfrak{M}[t; \Phi(t, r)].$$

If in Theorem 6.3 we let each C_n^* be identical with C^*, all the hypotheses of that theorem are satisfied, and the limit curve C_0^* is identical with C^*. So we need only show that the representation of C_0^* arrived at in the proof of Theorem 6.3 has the desired properties. Throughout the proof, t had the range $[0, 1]$, so (7.3) is satisfied. Property (7.4) is trivial, because of Lemma 4.1. The existence of the derivative in (7.6) was stated in (6.18). The proof of the equation (7.6) is the same as that of (6.23), with trivial notational changes.

Let $\Phi(r) = |r|$. If $0 \leqq \alpha \leqq \beta \leqq 1$, we have

$$\lim_{n\to\infty}\{\varphi_n(\beta) - \varphi_n(\alpha)\} = \varphi_0(\beta) - \varphi_0(\alpha).$$

That is,

$$\int_\alpha^\beta \mathfrak{M}_n[t; |r|]\,dt \to \int_\alpha^\beta \mathfrak{M}_0[t; |r|]\,dt.$$

But by (6.1)

$$\lim_{n\to\infty}\int_\alpha^\beta \mathfrak{M}_n[t; |r|]\,dt \leqq (\beta - \alpha)\mathfrak{L}(C_0^*).$$

Hence

$$\int_\alpha^\beta \mathfrak{M}_0[t; |r|]\,dt \leqq (\beta - \alpha)\mathfrak{L}(C_0^*). \tag{7.7}$$

In particular, (7.7) is an equality for $\alpha = 0$, $\beta = 1$. Hence by addition it must also be an equality for $\alpha = 0$, $\beta = t$ and for $\alpha = t$, $\beta = 1$. Thus

$$\int_0^t \mathfrak{M}_0[t; |r|]\,dt = t\mathfrak{L}(C_0^*).$$

By differentiation, with (7.6), this yields (7.5).

Another closely related special representation is described in the following lemma.

LEMMA 7.2. *If* C^* *is a rectifiable generalized curve, it has a representation* (7.2) *with properties* (7.3) *and* (7.6), *and with the additional properties*:
(7.8) *for all* t *in* M, *the value of* $\mathfrak{M}[t; \Phi]$ *is determined by the values of* Φ *on the set* $|r| = \mathfrak{L}(C^*)$;
(7.9) *for all* t *in* M, $\mathfrak{M}[t; 1] = 1$.

Let us first consider the representation

$$C^*: [y(t), \mathfrak{M}_0[t; \Phi], M] \tag{7.10}$$

which has the properties (7.3) to (7.6). If $\mathcal{L}(C^*) = 0$, by Lemma 2.2 we must have $\mathfrak{M}_0[t; \Phi] = 0$ for all t in M and all Φ. If we define

$$\mathfrak{M}[t; \Phi] = \Phi(0),$$

we readily verify that

$$[y(t), \mathfrak{M}[t; \Phi], M] \tag{7.11}$$

is another representation of C^* and has the desired properties. If $\mathcal{L}(C^*) > 0$, by Lemma 4.1 we can replace the representation (7.10) of C^* by another representation (7.11) which satisfies (7.8). From the formula (4.5) it is evident that property (7.6) persists. To verify (7.9), we observe that on the sphere $|r| = \mathcal{L}(C^*)$ the functions 1 and $|r|/\mathcal{L}(C^*)$ coincide; hence, by (7.8), (7.5) and (4.5), for all t in M we have

$$\mathfrak{M}[t; 1] = \mathfrak{M}[t; |r|]/\mathcal{L}(C^*) = 1.$$

Following L. C. Young, we shall call the representation of C^* described in Lemma 7.2 the *standard representation* of C^*, and we shall call t the *standard parameter*. Because of (7.9), in the *standard representation* of C^* we can refer to the linear functional $\mathfrak{M}[t; \Phi]$ as a *linear mean*.

8. **Existence theorems.** The existence theorems which we wish to establish in this section state that subject to suitable hypotheses, a class K of generalized curves contains a generalized curve C_0^* such that[8] $\mathcal{F}(C_0^*) \leqq \mathcal{F}(C^*)$ for all generalized curves C^* of the class K.

Let μ denote the greatest lower bound of the integral $\mathcal{F}(C^*)$ on the class K. A *minimizing sequence* is a sequence of curves C_n^* of the class K such that $\mathcal{F}(C_n^*)$ tends to μ. We shall assume that the following condition holds.

(8.1) *There exists a minimizing sequence $\{C_n^*\}$ such that for some finite number N the tracks of all the curves C_n^* lie in the sphere $|y| \leqq N$ and all the curves C_n^* are of length at most N.*

This condition will be discussed in §9.

It is now easy to establish an existence theorem.

THEOREM 8.1. *Let K be a closed class of generalized curves. Let $F(y, r)$ be defined and continuous and positively homogeneous of degree 1 in r for all r and for all y in a closed set E containing the tracks of all curves C^* of K. Let hypothesis (8.1) hold. Then if the class K is not empty, it contains a curve C_0^* such that $\mathcal{F}(C_0^*) \leqq \mathcal{F}(C^*)$ for all curves C^* of the class K.*

As we saw in §3, we may suppose $F(y, r)$ to be defined and continuous for all (y, r). Let $\{C_n^*\}$ be the minimizing sequence of hypothesis (8.1). By Theorem 6.3 there is a subsequence—we suppose it to be the whole sequence—

[8] Compare L. C. Young, loc. cit. (footnote 2), p. 233.

which converges to a limit curve C_0^*. Since K is closed, this curve is in the class K. By Theorem 6.1 we have

$$\mathcal{F}(C_0^*) = \lim_{n\to\infty} \mathcal{F}(C_n^*) = \mu;$$

and since μ is the greatest lower bound of $\mathcal{F}(C^*)$ on the class K, our theorem is established.

This theorem is easily extended to one of Bolza type.

THEOREM 8.2. *Let the hypotheses of Theorem 8.1 be satisfied. Let $g(z^1, \cdots, z^{2\nu})$ be defined and continuous on a closed point set P in 2ν-dimensional space. For each curve*

$$C^*: [y(t), \mathfrak{M}[t; \Phi(r)], M]$$

of the class K let the point

$$(y^1(a), \cdots, y^\nu(a), y^1(b), \cdots, y^\nu(b))$$

belong to P; we here denote the interval $\bar{M}$ by $[a, b]$. Then if the class K is not empty, it contains a curve C^* for which the functional

$$\mathcal{F}(C^*) + g(y^1(a), \cdots, y^\nu(b)) \tag{8.2}$$

assumes its minimum value.

The coördinates $y^i(a)$ and $y^i(b)$ are continuous functionals of the curve C^*, by the definition (4.1). Hence $g(y^1(a), \cdots, y^\nu(b))$ is also a continuous functional of C^*. The same is then true of the sum (8.2), and the proof of Theorem 8.1 can be applied at once.

9. **Discussion of the existence theorems.** It is easy to impose conditions on the integrand $F(y, r)$ which will insure the fulfilling of condition (8.1). For instance, we may require that the set E be bounded and that $F(y, r)$ be positive for y in E and $|r| \neq 0$. Or, E being unbounded, we may require that there exist a positive number k such that the inequality

$$F(y, r) \geqq K\,|\,r\,|\,(1 + |\,y\,|)^{-1}$$

holds for all y in E and all r. Moreover, the theorems of Hahn[9] and Tonelli[10] can be extended without difficulty to generalized curves.

Examples of closed classes K can be constructed by restricting the track alone. For instance, K may consist of the curves C^* whose initial points lie in a given closed set A_1 and whose terminal points lie in a given set A_2. Or we may require the point $(y^1(a), \cdots, y^\nu(b))$ to lie in a given closed set P in 2ν-dimen-

[9] H. Hahn, *Über ein Existenztheorem der Variationsrechnung*, Sitzungsberichte Akad. Wiss. Wien, Mathem. Naturwiss. Kl., Abt. II-a, vol. 134(1925), pp. 437–447.

[10] L. Tonelli, *Sull'esistenza del minimo in problemi di calcolo delle variazioni*, Annali della Reale Scuola Normale Superiore di Pisa, ser. II, vol. 1(1932), pp. 89–99.

sional space, as in Theorem 8.2. Or we may require that the track of C^* shall have a point in common with a given closed set.

This is familiar. But the use of generalized curves permits the imposition of the condition that an integral

$$\mathcal{G}(C^*) = \int_a^b \mathfrak{M}[t; G(y(t), r)]\, dt \tag{9.1}$$

shall have an assigned value γ. For by Theorem 6.1, if $\mathcal{G}(C_n^*) = \gamma$ and C_n^* converges to C_0^*, then $\mathcal{G}(C_0^*) = \gamma$.

An *ordinary* curve $y = y(t)$ satisfies the differential equation $\varphi(y, y') = 0$ if the equation $\varphi(y(t), y'(t)) = 0$ holds for almost all t. This is equivalent to demanding that the integral of $| \varphi(y(t), \dot{y}(t)) |$ shall vanish. We adopt the last formulation as a basis for extending to generalized curves. If C^*: $[y(t), \mathfrak{M}[t; \Phi], M]$ is a generalized curve and $\varphi(y, r)$ is a parametric integrand, we shall say that C^* satisfies the differential equation

$$\varphi(y, r) = 0 \tag{9.2}$$

if the equation

$$\int_M \mathfrak{M}[t; | \varphi(y(t), r) |]\, dt = 0 \tag{9.3}$$

is satisfied.

Since a differential equation (9.2) is expressible as an isoperimetric condition (9.3), the curves satisfying such a requirement form a closed set.

Again, a differential inequality

$$\psi(y, r) \geqq 0 \tag{9.4}$$

can be written as a differential equation

$$\varphi(y, r) \equiv \min\, [0, \psi(y, r)] = 0. \tag{9.5}$$

Hence the class of generalized curves satisfying a differential inequality (9.4) is a closed class.

Finally, the product of any aggregate of closed sets is closed. It follows that if K is a class consisting of all curves lying in a closed set E, having a track which satisfies any finite or infinite collection of conditions of the type discussed in the second paragraph of this section, and satisfying any finite or infinite collection of isoperimetric conditions, differential equations and differential inequalities, then K is closed.

10. **Vectors carried by a generalized curve.** The integral formulation (9.3) of the differential equation (9.2), although effective in §9, is not convenient for later use. In (1.2) we notice that the two vectors (1, 1) and (1, −1) are distinguished by the fact that they, and they alone, are used in computing the mean. We shall say that these vectors are *carried* by the curve. In general,

if C^*: $[y(t), \mathfrak{M}[t; \Phi], M]$ is a generalized curve, and t_0 is a point of the set M, we say that the vector r_0 is *carried by* C^* *at* $y(t_0)$ if for every non-negative $\Phi(r)$ such that $\Phi(r_0) > 0$, the inequality

$$\mathfrak{M}[t_0 ; \Phi(r)] > 0 \tag{10.1}$$

is satisfied. It is easily seen that the class of vectors carried at $y(t_0)$ is a closed set, and by (2.8b) it is bounded.

We now prove a lemma.

LEMMA 10.1. *For each t_0 in M, the value of $\mathfrak{M}[t_0 ; \Phi]$ is determined by the values of Φ on the set K of vectors carried by C^* at $y(t_0)$.*

If r_0 is a vector not carried at $y(t_0)$, there is a non-negative function $\Phi(r)$ such that $\Phi(r_0)$ is positive and $\mathfrak{M}[t_0 ; \Phi] = 0$. The function Φ remains positive on a neighborhood of r_0. A denumerable set of these neighborhoods covers the complement CK of the set K. Thus we have a sequence $\{\Phi_n(r)\}$ of non-negative continuous functions such that

$$\mathfrak{M}[t_0 ; \Phi_n] = 0, \tag{10.2}$$

while at each point r_0 of CK at least one Φ_n is positive. We may suppose that

$$\text{l.u.b. } \Phi_n(r) \leqq 2^{-n}, \tag{10.3}$$

since otherwise we can replace Φ_n by $\min [\Phi_n(r), 2^{-n}]$. It follows that the series

$$\Phi_0(r) = \sum \Phi_n(r)$$

is uniformly convergent. By (10.2) and Lemma 2.2, this implies

$$\mathfrak{M}[t_0 ; \Phi_0] = 0. \tag{10.4}$$

The function Φ_0 is non-negative, and is positive on CK.

By (2.8b), the value of $\mathfrak{M}[t_0 ; \Phi]$ depends only on the values of Φ on some sphere $|r| \leqq H$. Let $\Phi(r)$ be a continuous function which vanishes on K. It is easily seen that if ϵ is a positive number, there is a positive number N such that

$$N\Phi_0(r) + \epsilon > |\Phi(r)|$$

for all r such that $|r| \leqq H$. By (10.4),

$$\begin{aligned} |\mathfrak{M}[t_0 ; \Phi]| &\leqq \mathfrak{M}[t_0 ; |\Phi|] \\ &\leqq \epsilon\mathfrak{M}[t_0 ; 1]. \end{aligned}$$

Since ϵ is arbitrary, the left member of this inequality must vanish, and the lemma is established.

It is now easy to prove the following lemma.

LEMMA 10.2. *The generalized curve C^*: $[y(t), \mathfrak{M}[t; \Phi], M]$ satisfies the differential equation $\varphi(y, r) = 0$ if and only if for almost all t in M the equation*

$$\varphi(y(t), r) = 0 \tag{10.5}$$

is satisfied for all vectors r carried at $y(t)$.

If (9.3) holds, for almost all t we have

$$\mathfrak{M}[t; \mid \varphi(y(t), r) \mid] = 0. \tag{10.6}$$

By definition, no vector at which $\mid \varphi(y(t), r) \mid > 0$ is carried at $y(t)$, so (10.5) holds for all vectors carried. Conversely, if (10.5) holds for all vectors r carried at $y(t)$, from Lemma 10.1 we deduce equation (10.6). Since this holds for almost all t, equation (9.3) is satisfied.

Another easy consequence of Lemma 10.1 is the following.

LEMMA 10.3. *Let C^*: $[y(t), \mathfrak{M}[t; \Phi], M]$ be a generalized curve. Let t be a point of M, and $K(t)$ the set of vectors carried by C^* at $y(t)$. If $\Phi(r)$ is a continuous function which is non-negative on $K(t)$, either it vanishes for all r in $K(t)$ or else*

$$\mathfrak{M}[t; \Phi] > 0. \tag{10.7}$$

The functions $\Phi(r)$ and

$$\Phi_0(r) \equiv \max [0, \Phi(r)]$$

coincide on $K(t)$ by hypothesis, so by Lemma 10.1 we have $\mathfrak{M}[t; \Phi] = \mathfrak{M}[t; \Phi_0] \geqq 0$. If equality holds, by definition of vectors carried no vector for which $\Phi_0(r)$ is positive can be carried. That is, $\Phi(r)$ is non-positive on $K(t)$, and the lemma is established.

11. **Some theorems on the existence of a minimizing ordinary curve.** Let us say that a class K of rectifiable (ordinary) curves is *quasi-closed* if for every N the subclass consisting of these curves of K with length at most N is a closed class, in the Fréchet topology. In this section we shall consider the problem of minimizing an integral

$$\mathcal{F}(C) = \int_a^b F(y, \dot{y})\, dt \tag{11.1}$$

in the subclass K_0 of K consisting of those curves C: $y = y(t)$ $(a \leqq t \leqq b)$ of K which satisfy conditions of the following types.

(a) Isoperimetric conditions of the form

$$\int_a^b L_\alpha(y, \dot{y})\, dt = \lambda_\alpha. \tag{11.2}$$

(b) Isoperimetric inequalities of the form

$$\int \mathcal{G}_\beta(y, \dot{y})\, dt \leqq \gamma_\beta. \tag{11.3}$$

(c) Differential equations

$$\varphi_\gamma(y, \dot{y}) = 0. \tag{11.4}$$

(d) Differential inequalities

$$\psi_\delta(y, \dot{y}) \leqq 0. \tag{11.5}$$

Subject to drastic requirements on the functions involved, we shall show the existence of a curve which minimizes $\mathcal{F}(C)$ in the class of curves satisfying the conditions listed.

To begin with, we note that if for each y in a set E the functions $\varphi_\gamma(y, r)$ are linear on r and the functions $\psi_\delta(y, r)$ convex in r, we can readily show that the aggregate $Q(y)$ of vectors r such that

$$\varphi_\gamma(y, r) = 0 \quad \text{and} \quad \psi_\delta(y, r) \leqq 0 \tag{11.6}$$

is convex for each y in E. These conditions on φ_γ and ψ_δ are sufficient, but not necessary, for the convexity of $Q(y)$. We now suppose that our problem satisfies the following requirements.

(11.7) E is a closed set.

(11.8) K is a quasi-closed class of rectifiable curves lying in E.

(11.9) The functions F, L_α, G_β, φ_γ, ψ_δ are defined and continuous for all y in E and all r, and are positively homogeneous of degree 1 in r.

(11.10) For each y in E, the aggregate $Q(y)$ of solutions of (11.6) is convex.

(11.11) For each y in E, the functions $F(y, r)$ and $G_\beta(y, r)$ are convex on $Q(y)$ and the $L_\alpha(y, r)$ are linear on $Q(y)$.

(11.12) At least one curve of K satisfies conditions (a), (b), (c), (d); that is, the subclass K_0 is not empty.

(11.13) There exists a minimizing sequence $\{C_n\}$ for $\mathcal{F}(C)$ in the class K such that for some finite N all the curves C_n lie in the sphere $|y| \leqq N$ and have length at most N.

We can prove the following theorem.

Theorem 11.1. *If hypotheses* (11.7) *to* (11.13) *are satisfied, the subclass* K_0 *contains a curve* C_0 *which gives* $\mathcal{F}(C)$ *its least value* K_0.

Let m be the greatest lower bound of the values of $\mathcal{F}(C)$ on K_0. By hypothesis (11.13) there is a sequence $\{C_n\}$ of curves of K lying in the sphere $|y| \leqq N$ and having length at most N, and such that $\mathcal{F}(C_n)$ tends to m. By Theorem 6.3, we can choose a subsequence (we may suppose it to be the whole sequence) which converges to a generalized curve. We suppose that

$$C_0^* : [y_0(t), \mathfrak{M}_0[t; \Phi(r)], M] \tag{11.14}$$

is the standard representation of C_0^*. From hypotheses (11.7) and (11.8) it follows that the ordinary curve

$$C_0 : y = y_0(t) \qquad (0 \leqq t \leqq 1) \tag{11.15}$$

(the track of C_0^*) lies in E and belongs to K. As we saw in §9, the generalized curve C_0^* will satisfy the conditions

$$\int_a^b \mathfrak{M}_0[t; F(y_0(t), r)]\, dt = m, \tag{11.16}$$

$$\int_a^b \mathfrak{M}_0[t; L_\alpha(y_0(t), r)]\, dt = \gamma_\alpha, \tag{11.17}$$

$$\int_a^b \mathfrak{M}_0[t; G_\beta(y_0(t), r)]\, dt \leqq \gamma_\beta, \tag{11.18}$$

$$\varphi_\gamma(y_0(t), r) = 0, \tag{11.19}$$

$$\psi_\delta(y_0(t), r) \leqq 0, \tag{11.20}$$

the last two being understood to hold for almost all t and for all r carried at $y_0(t)$. These last two conditions imply that for all t in a set of measure $b - a$, all vectors r carried at $y_0(t)$ lie in the convex set $Q(y_0(t))$.

By Jensen's inequality[11] in the geometric form the mean

$$(\mathfrak{M}_0[t; r^1], \cdots, \mathfrak{M}_0[t; r^\nu])$$

is in $Q(y_0(t))$. That is, by (2.8c), $\dot{y}_0(t)$ is in $Q(y_0(t))$, and conditions (c) and (d) are satisfied by C_0.

From the linearity of $L_\alpha(y, r)$ we have

$$L_\alpha(y_0(t), \dot{y}_0(t)) = L_\alpha(y_0(t), \mathfrak{M}_0[t; r]) = \mathfrak{M}_0[t; L_\alpha(y_0(t), r)] \tag{11.21}$$

for all t in M. On integration this yields

$$\int_a^b L_\alpha(y_0, \dot{y}_0)\, dt = \int_a^b \mathfrak{M}_0[t; L_\alpha(y_0(t), r)]\, dt,$$

so that C satisfies condition (a).

By Jensen's inequality and hypothesis (11.11) we find

$$\mathfrak{M}_0[t; G_\alpha(y_0(t), r)] \geqq G_\alpha(y_0(t), \dot{y}_0(t)), \tag{11.22}$$

$$\mathfrak{M}_0[t; F(y_0(t), r)] \geqq F(y_0(t), \dot{y}_0(t)), \tag{11.23}$$

[11] See, e.g., E. J. McShane, *Jensen's inequality*, Bull. Amer. Math. Soc., vol. 43(1937), pp. 521–527.

I wish to take this opportunity to publish the answer to a question raised in that paper. I there stated that I did not know whether the hypothesis that $f(x) = c$ except on a negligible set was sufficient to insure that $\varphi(Mf) = M\varphi(f)$. Professor Bohnenblust points out that the answer is in the negative. Let L be the class of all functions $f(x)$ defined for all real x and having the limits $f(1-)$, $f(-1+)$ defined, and let $Mf = \frac{1}{2}\{f(-1+) + f(1-)\}$. The whole real axis is a negligible set, for the function

$$f(x) = 1 - |x|, \qquad |x| < 1,$$
$$f(x) = 1, \qquad |x| \geqq 1,$$

is everywhere positive and yet $Mf = 0$. Take $\varphi(z) = z^2$, $f(x) = x$. Then $f(x) = 0$ except on a negligible set (all x!), but $M\varphi(f) = M(x^2) = 1 \neq 0 = \varphi(Mf)$.

for all t in M. Integration in (11.22) shows that C_0 satisfies condition (b), and this completes the proof that C is in K. By definition of m, we have

$$\mathcal{F}(C_0) \geqq m. \tag{11.24}$$

On the other hand, integration in (11.23) yields $\mathcal{F}(C_0) \leqq m$. This, with (11.24), proves that $\mathcal{F}(C_0) = m$, so that C_0 is the curve sought.

Theorem 11.1 is essentially equivalent to several theorems in the literature.[12] Up to this stage in our study, the use of generalized curves in proving existence theorems for ordinary curves has only yielded a new proof of a known result. In a later paper we shall see that it is possible to establish theorems of a nature widely different from that of Theorem 11.1.

Appendix

12. **Metrization of the space of generalized curves.** Although it is not essential for our purposes, it is interesting to notice that the space of rectifiable generalized curves, with the topology (4.1), is metrizable. Let us first select a special class of integrands:

(12.1) $(L, 1)$ *is the class of all parametric integrands* $F(y, r)$ *which on the point-set* [*all* y, r *such that* $|r| = 1$] *are at most* 1 *in absolute value and satisfy a Lipschitz condition of constant* 1.

Our distance definition is the following

(12.2) *For each pair* C_1^*, C_2^* *of generalized curves the distance* $\| C_1^*, C_2^* \|$ *is the sum of* (i) *the Fréchet distance* $\| y_1(x), y_2(x) \|$ *between the tracks of the curves, and* (ii) *the least upper bound of* $| \mathcal{F}(C_1^*) - \mathcal{F}(C_2^*) |$ *for all integrands* $F(y, r)$ *in the class* $(L, 1)$.

We verify at once that this distance is symmetric and non-negative, and the triangle inequality is easy to establish. In order that $\| C_1^*, C_2^* \| = 0$, it is clearly sufficient (by (4.2)) that $C_1^* = C_2^*$. It remains to show that this is also necessary, and to prove that for every sequence $C_0^*, C_1^*, \cdots$, the equation $\lim_{n\to\infty} C_n^* = C_0^*$ holds if and only if $\lim_{n\to\infty} \| C_n^*, C_0^* \| = 0$.

Suppose that $\lim_{n\to\infty} \| C_n^*, C_0^* \| = 0$. Then by definition (12.2) condition (4.1a) is satisfied. Also, since the integrand $|r|$ is in $(L, 1)$, we have

$$\lim_{n\to\infty} | \mathcal{L}(C_n^*) - \mathcal{L}(C_0^*) | = 0. \tag{12.3}$$

Let $F(y, r)$ be a parametric integrand. If Δ is an interval large enough to contain the tracks of all the C_n^* and of C_0^*, and ϵ is an arbitrary positive number, we can find a polynomial $P(y, r)$ such that

$$| P(y, r) - F(y, r) | \leqq \epsilon \tag{12.4}$$

[12] See, e.g., L. M. Graves, *On the existence of the absolute minimum in problems of Lagrange*, Bull. Amer. Math. Soc., vol. 39(1933), pp. 101–104.

for all y in Δ and all r such that $|r| \leqq 1$. On this set $P(y, r)$ is Lipschitzian, so it can be extended to be defined and Lipschitzian[13] on all of (y, r)-space. We now render it homogeneous by the formula (4.3) with $k = 1$; we thus obtain an integrand $F_1(y, r) = P_h(y, r)$ which is Lipschitzian on the set of (y, r) such that $|r| = 1$, and satisfies the inequality

$$|F_1(y, r) - F(y, r)| \leqq \epsilon |r| \tag{12.5}$$

for all y in Δ and all r. For a sufficiently small positive number γ the integrand γF_1 is in $(L, 1)$, so that by (12.2) the inequality

$$|\gamma\mathcal{F}_1(C_n^*) - \gamma\mathcal{F}_1(C_0^*)| < \gamma\epsilon$$

holds for all large n. This, with (12.5), shows that for all large n the inequality

$$|\mathcal{F}(C_n^*) - \mathcal{F}(C_0^*)| < \epsilon\mathcal{L}(C_n^*) + \epsilon\mathcal{L}(C_0^*) + \epsilon$$

holds. But, ϵ being arbitrary, this and (12.3) imply that $\lim \mathcal{F}(C_n^*) = \mathcal{F}(C_0^*)$. Hence condition (4.1b) is also satisfied, and C_n^* tends to C_0^*.

Incidentally, if $\|C_1^*, C_2^*\| = 0$, then for the sequence C_n^* $(n = 2, 3, \cdots)$ with $C_n^* = C_2^*$ for all n we obviously have $\lim \|C_1^*, C_n^*\| = 0$. Hence by the preceding proof we know that $C_n^* \to C_1^*$, and this is only possible if $C_2^* = C_1^*$.

Conversely, suppose that $\|C_n^*, C_0^*\|$ does not tend to zero; we must show that C_n^* does not approach C_0^*. We select a subsequence (retaining the notation for the original sequence) such that $\|C_n^*, C_0^*\|$ tends to a finite or infinite limit different from zero. If the Fréchet distance between the track of C_n^* and that of C_0^* does not tend to zero, then by (4.1) C_n^* does not tend to C_0^*. If the Fréchet distance does approach zero, there must exist a positive ϵ and (for each n) an integrand F_n of the class $(L, 1)$ such that

$$|\mathcal{F}_n(C_n^*) - \mathcal{F}_n(C_0^*)| \geqq \epsilon. \tag{12.6}$$

Let $\bar{U}$ be a closed neighborhood of the track of C_0^*. By Ascoli's theorem there exists a subsequence of the F_n (we retain the original notation) which converges to a limit function $F_0(y, r)$, uniformly on the set $[y \text{ in } \bar{U}, |r| = 1]$. Then for every positive number γ the inequality

$$|F_n(y, r) - F_0(y, r)| \leqq \gamma |r|$$

holds for all y in $\bar{U}$ and all r. Hence

$$|\mathcal{F}_n(C_n^*) - \mathcal{F}_n(C_0^*)| - |\mathcal{F}_0(C_n^*) - \mathcal{F}_0(C_0^*)| \leqq \gamma\{\mathcal{L}(C_n^*) + \mathcal{L}(C^*)\}. \tag{12.7}$$

[13] E. J. McShane, *Extension of range of functions*, Bull. Amer. Math. Soc., vol. 40(1934), pp. 837–842.

If equation (12.3) is false, C_n^* does not tend to C_0^*, since condition (4.1b) fails; and then our proof is complete. If (12.3) is valid, the sum in the braces in (12.7) is bounded, and by choosing γ small enough we find that (12.7) and (12.6) imply that

$$| \mathcal{F}_0(C_n^*) - \mathcal{F}_0(C_0^*) | > \tfrac{1}{2}\epsilon$$

for all large n. Hence C_n^* fails to approach C_0^*. This completes the proof that the metric (12.2) yields the topology (4.1).

THE UNIVERSITY OF VIRGINIA.

THE APPROXIMATION OF FUNCTIONS SATISFYING A LINEAR PARTIAL DIFFERENTIAL EQUATION

By Stefan Bergman

1. **Statement of the problem.** The determination of the solution of a linear differential equation satisfying given boundary data is a classical problem in the theory of these equations. For instance, in the case of elliptic equations[1] values on the boundary of a domain may be given, and it is necessary to search for the function which satisfies the equation inside the domain and which assumes the given values on the boundary (first boundary problem). If, moreover, these values are equal to zero, it is necessary in certain cases to determine a parameter in the equation for which there exists a non-identically vanishing solution which is zero on the boundary (characteristic value problem). In the case of hyperbolic equations, values of the function and of its first derivatives are known on the initial line, which is cut by every characteristic line at one point only; from these data it is necessary to determine the solution in a certain domain (Cauchy's problem). In addition to the problem of the proof of the existence of solutions of this kind, there is the problem of finding a method of calculating numerically the desired solution from the given data.[2]

In the papers [2, 5, 6, 7][3] stress has been laid on a connection between functions $F(x, y)$ satisfying a linear partial differential equation in two real variables x, y and certain classes of functions $f(z)$ of one complex or real variable. In fact, given an equation $\mathbf{L}(U) = 0$, it is possible to find a linear operation which transforms an arbitrary function $f(z)$ of a certain class into a particular solution of $\mathbf{L}(U) = 0$. As a well-known special case, we obtain a harmonic function (i.e., a function satisfying Laplace's differential equation) if we take the real part of any analytic function $f(z)$.

This relation enables us to transfer certain theorems of the theory of functions of one variable to functions which satisfy a linear differential equation. In particular, we approximate arbitrary solutions of a differential equation by linear combinations of a system of particular solutions of the *same* equation; and in this manner deal with the problems mentioned above.

The following two problems arise (both of which will be solved in the present paper):

(1) Given a differential equation $\mathbf{L}$, to give a procedure for calculating a

Received March 17, 1939.

[1] Elliptic (hyperbolic) equation means linear differential equation of elliptic (hyperbolic) type.

[2] The proof of the existence of a solution often does not enable one to calculate this solution; e.g., if we use the method of choice. Often, also, the determination of the solution along the lines indicated in the proof leads to enormous numerical calculations.

[3] Numbers in brackets refer to the bibliography.

system of particular solutions P_n ($n = 1, 2, \cdots$) of **L** with the property that, if U is an arbitrary solution of **L** regular in a domain $\mathfrak{F}^2$, then,[4] for every $\mathfrak{T}^2 \subset \mathfrak{F}^2$ and every $\epsilon > 0$, there exists an expression $W = \sum_{\nu=1}^{n} a_\nu P_\nu$ such that $| U - W | \leqq \epsilon$ in $\bar{\mathfrak{T}}^2$.

(2) To give a method for calculating the coefficients $A_\nu^{(n)}$ in the linear form $W_n = \sum_{\nu=1}^{n} A_\nu^{(n)} P_\nu$ such that W_n converges for $n \to \infty$ to the desired solution of **L**. In the case of characteristic value problems, our procedure must permit calculation of the parameter in the equation for which there exists a solution U, different from $U \equiv 0$, vanishing on the boundary.

Notations. In this paper we designate by $\mathfrak{F}^2$ a convex domain of the plane, containing the origin; by $\mathfrak{f}^1$ its boundary, a curve which everywhere possesses a continuously changing curvature; in the usual manner we designate $\mathfrak{F}^2 + \mathfrak{f}^1$ by $\bar{\mathfrak{F}}^2$. By s we denote the length of the arcs of $\mathfrak{f}^1$ beginning with a certain fixed point—for instance, the intersection of $\mathfrak{f}^1$ with the positive x-axis.

Analytic functions of one complex variable z or u will be denoted by *small* letters and functions of z, $\bar{z}$ or z, $\bar{z}$; ζ, $\bar{\zeta}$ by *capital* letters; e.g., $f(z)$ is an analytic function of z; $F(z)$ a function of z, $\bar{z}$ (in extended form, $F(z, \bar{z})$); and $K(z, \zeta)$ a function of $z, \bar{z}; \zeta, \bar{\zeta}$ (in extended form, $K(z, \bar{z}; \zeta, \bar{\zeta})$). If the function depends in addition on a parameter δ or t, we shall write it as $K(z, \zeta; \delta)$ or $F(z; t)$ in place of the extended form with all variables given explicitly.

ϵ will denote a positive number, arbitrarily small, or alternatively will denote the relation "is contained in".

2. **Statement of some preliminary results.** Before dealing with problems (1) and (2), we point out some preliminary results to be applied to the treatment of problem (1). Define

$$\mathbf{G}(U) \equiv U_{xy} + A(x, y)U_x + B(x, y)U_y + C(x, y)U = 0, \tag{2.1}$$

where $U_x = \partial U/\partial x$, $U_y = \partial U/\partial y$, and x, y are Cartesian coördinates in the plane. Let $\mathfrak{S}^1$ represent a differentiable open curve in the plane of a complex variable t which is contained within the circle $| t | \leqq 1$ and which touches the real axis at $t = \pm 1$.

THEOREM 1. *Suppose A, B, and C are functions with continuous first derivatives, and that* $\mathrm{E}(x, y, t)$ *is a function with continuous second derivatives satisfying the equation*

$$-(1 - t^2)\mathrm{E}^*_{yt} + \frac{1}{t}\,\mathrm{E}^*_y - 2tx[\mathrm{E}^*_{xy} + D\mathrm{E}^*_y + F\mathrm{E}^*] = 0, \tag{2.2}$$

[4] The superscript number indicates the dimension of the manifold. A bar over $\mathfrak{T}^2$ and $\mathfrak{F}^2$ indicates that $\bar{\mathfrak{T}}^2$ and $\bar{\mathfrak{F}}^2$ are closed.

where

$$D \equiv D(x, y) = n_x(x) - \int_0^y A(x, y)\, dy + B(x, y),$$

$$F \equiv F(x, y) = -A_x(x, y) - A(x, y)B(x, y) + C(x, y),$$

and $n(x)$ is an arbitrary function of a real variable x possessing continuous second derivatives. Then, if $f(u)$ is an arbitrary function of one real variable u with continuous second derivatives, we have

$$U(x, y) = \int_{\mathfrak{S}^1} \mathbf{E}(x, y, t) f\left(x \cdot \frac{1 - t^2}{2}\right)(1 - t^2)^{-1/2}\, dt, \tag{2.3}$$

where

$$\mathbf{E}(x, y, t) = \mathbf{E}^*(x, y, t) \exp\left[n(x) - \int_0^y A(x, y)\, dy\right] \qquad (\exp a \equiv e^a), \tag{2.4}$$

is a particular solution of (2.1). (See [6], p. 1172.)

In order to prove this, it is sufficient to show that the expression

$$V(x, y) = \int_{\mathfrak{S}^1} \mathbf{E}^*(x, y, t)\, f\left(x \cdot \frac{1 - t^2}{2}\right)(1 - t^2)^{-1/2}\, dt$$

satisfies the equation

$$\mathbf{g}^*(V) \equiv V_{xy} + DV_y + FV = 0. \tag{2.5}$$

We have

$$V_y = \int_{\mathfrak{S}^1} \mathbf{E}_y^* \cdot f \cdot (1 - t^2)^{-1/2}\, dt, \tag{2.6a}$$

$$V_{xy} = \int_{\mathfrak{S}^1} \mathbf{E}_{xy}^* \cdot f \cdot (1 - t^2)^{-1/2}\, dt - \int_{\mathfrak{S}^1} \mathbf{E}_y^* (2xt)^{-1}(1 - t^2)^{1/2} f_t\, dt, \tag{2.6b}$$

since $f_x = -f_t \cdot [(1 - t^2)/(2xt)]$. Integrating by parts in the second integral in (2.6b) and substituting the resulting expression, together with (2.6b), in (2.5), we find by (2.2) that $\mathbf{g}^*(V) = 0$.

If $\mathbf{E}^*$ can be represented in the form

$$\mathbf{E}^* = C_0 + txy\mathbf{E}_1^*(x, y, t), \tag{2.7}$$

where $\mathbf{E}_1^*$ is a function possessing continuous derivatives of the first order and C_0 is a constant, we say that $\mathbf{E}$ is a *generating function of the first kind*.

In the following, we shall consider equations (2.1) whose coefficients A, B, C are analytic in a sufficiently large four-dimensional domain $|x| \leq r$, $|y| \leq r$ (x and y *here* denote two complex variables). In this case, the following theorem has been established ([6], pp. 1173–1177).

THEOREM 2. *For every equation* $\mathbf{G}(U) = 0$ (see (2.1)) *there exist two generating functions* $\mathbf{E}_k = H_k\mathbf{E}_k^*$ $(k = 1, 2)$,

$$H_1 = \exp\left[n_1(x) - \int_0^y A(x, y)\,dy\right], \qquad H_2 = \exp\left[n_2(y) - \int_0^x B(x, y)\,dx\right],$$

of the first kind such that

$$U(x, y) = \int_{\mathfrak{S}^1}\left[\mathbf{E}_1(x, y, t)f\left(x\cdot\frac{1-t^2}{2}\right) + \mathbf{E}_2(x, y, t)g\left(y\cdot\frac{1-t^2}{2}\right)\right](1-t^2)^{-\frac{1}{2}}\,dt \tag{2.8}$$

is a particular solution of $\mathbf{G}(U) = 0$ *if* $f(u)$ *and* $g(u)$ *are arbitrary functions of* u *possessing continuous second derivatives. Conversely, every solution* U *of* $\mathbf{G}(U) = 0$ *may be written in the form* (2.8) *in the neighborhood of* $x = y = 0$, *where*

$$\begin{aligned} f(u) &= \frac{2}{\pi\mathbf{E}_1(0, 0, 0)}\int_0^{\frac{1}{2}\pi} u\sin\theta\,\frac{dW_1(2u\sin^2\theta, 0)}{d(u\sin^2\theta)}\,d\theta + \frac{W_1(0, 0)}{\pi\mathbf{E}_1(0, 0, 0)}, \\ g(u) &= \frac{2}{\pi\mathbf{E}_2(0, 0, 0)}\int_0^{\frac{1}{2}\pi} u\sin\theta\,\frac{dW_2(0, 2u\sin^2\theta)}{d(u\sin^2\theta)}\,d\theta, \end{aligned} \tag{2.9}$$

and where

$$W_1(x, y) = \frac{U(x, y)}{H_1(x, y)}, \qquad W_2(x, y) = \frac{U(x, y) - 2W_1(0, 0)H_1(x, y)}{H_2(x, y)}.$$

We suppose from now on that $\mathbf{E}_1$ and $\mathbf{E}_2$ are generating functions of the first kind and that $\mathfrak{S}^1$ is the interval $\{-1, 1\}$ of the real axis. In the case of generating functions of the first kind, $\mathfrak{S}^1$ may contain the point $t = 0$. In [6] (pp. 1173–1176) $\mathbf{E}_1$ and $\mathbf{E}_2$ have been calculated in the form of infinite series.

For elliptic equations we find analogous results. Let $z = X + iY$, $\bar{z} = X - iY$, where X, Y again mean Cartesian coördinates in the plane. We understand now by U_z and $U_{\bar{z}}$ respectively

$$U_z = \frac{1}{2}\left(\frac{\partial U}{\partial X} - i\,\frac{\partial U}{\partial Y}\right), \qquad U_{\bar{z}} = \frac{1}{2}\left(\frac{\partial U}{\partial X} + i\,\frac{\partial U}{\partial Y}\right),$$

$$U_{z\bar{z}} = \frac{1}{4}\left(\frac{\partial^2 U}{\partial X^2} + \frac{\partial^2 U}{\partial Y^2}\right).$$

If we substitute z for x and $\bar{z}$ for y in (2.1) and suppose that $C(z)$ is a real function of z, $\bar{z}$ and that $A(z) = \overline{B(z)}$, then this equation takes the form

$$\mathbf{L}(U) \equiv U_{z\bar{z}} + 2\,\mathrm{Re}\,[A(z)U_z] + C(z)U = 0, \tag{2.10}$$

where "Re" means "the real part of". Therefore, for every solution U of $\mathbf{L}(U) = 0$ we obtain the representation

$$U(z) = \mathrm{Re}\left[\int_{-1}^{1}\mathbf{E}(z; t)f\left(z\cdot\frac{1-t^2}{2}\right)(1-t^2)^{-\frac{1}{2}}\,dt\right]. \tag{2.11}$$

As is well known, every solution $U(z)$, $\{z = X + iY\} \in \mathfrak{F}^2$, of an elliptic equation $\mathbf{L}(U) = 0$ is an analytic function of X and Y in $\mathfrak{F}^2$ ([10], p. 1320). Moreover, it is possible to prove that in this case the function $f(u)$ of (2.11) is an analytic function of $u = \frac{1}{2}(X + iY)(1 - t^2)$ in $\mathfrak{F}^2$ ([6], p. 1179).[5] These results permit us to construct systems of particular solutions which are complete in the sense indicated in problem (1), §1.

If $\mathbf{L}(U) = 0$ is an elliptic equation, the system

$$P_{2n-\alpha}(z) = \frac{\mathrm{Re}}{\mathrm{Im}}\left[2^{-\frac{1}{2}} z^n \int_{-1}^{1} \mathbf{E}(z; t)\left(\frac{1 - t^2}{2}\right)^{n-\frac{1}{2}} dt\right], \tag{2.12}$$

$$n = 0, \qquad \alpha = 0, \qquad n = 1, 2, \cdots, \begin{cases} \alpha = 1 \\ \alpha = 0 \end{cases}$$

is complete (see §3). In the case of a hyperbolic equation we can choose

$$\begin{aligned} \psi_{2n}(x, y) &= 2^{-\frac{1}{2}} x^n \int_{-1}^{1} \mathbf{E}_1(x, y, t)\left(\frac{1 - t^2}{2}\right)^{n-\frac{1}{2}} dt, \\ \psi_{2n+1}(x, y) &= 2^{-\frac{1}{2}} y^n \int_{-1}^{1} \mathbf{E}_2(x, y, t)\left(\frac{1 - t^2}{2}\right)^{n-\frac{1}{2}} dt, \end{aligned} \tag{2.13}$$

$$(n = 0, 1, 2, \cdots).$$

3. Proof that the system (2.12) is complete.

LEMMA 1. *Let*[6] *$U(z)$ satisfy the equation* (2.1) *and be regular in* $\bar{\mathfrak{F}}^2$. *Then*

[5] The fact that any solution $U(z)$ regular in $\mathfrak{F}^2$ is also regular in the domain $\mathfrak{Z}_1^4 \cdot \mathfrak{Z}_2^4$ of the space X_1, Y_1, X_2, Y_2, where $X = X_1 + iY_1$, $Y = X_2 + iY_2$, and

$$\mathfrak{Z}_k^4 = \mathrm{E}[X_1 + (-1)^k Y_2 = a,\ X_2 - (-1)^k Y_2 = b,\ (a + ib) \in \mathfrak{F}^2] \quad (k = 1, 2),$$

has an essential importance in the proof of the assertion made above. It can be proved not only in the way indicated in [6], p. 1178, but also with the aid of the formula

$$2\pi U(z) = \int_{\mathfrak{f}^1}\left[V(z, \zeta)\frac{\partial U(\zeta)}{\partial \mathrm{n}_\zeta} - U(\zeta)\frac{\partial V(z, \zeta)}{\partial \mathrm{n}_\zeta} + 8\,\mathrm{Re}\,\{A(\zeta)[\cos(\mathrm{n}, \xi) - i\cos(\mathrm{n}, \eta)]U(\zeta)V(z, \zeta)\}\right] ds,$$

where V is a solution of the adjoint equation

$$V_{\zeta\bar{\zeta}} - 2\,\mathrm{Re}\,[AV]_\zeta + CV = 0, \qquad \zeta = \xi + i\eta,$$

which in the neighborhood of $\zeta = z$ has the same behavior as $\log|z - \zeta|$ (see [11], p. 515). The existence of such a solution V, regular in a sufficiently large domain except for $\zeta = z$, is established in [9].

[6] We recall the notations of §1, which we shall follow here.

there exists for every $\epsilon > 0$ *a linear form in the* P_k (see (2.12)), $W(z) = \sum_{k=1}^{n} a_k P_k(z)$, *such that*

$$|U(z) - W(z)| \leqq \epsilon, \qquad z \in \bar{\mathfrak{F}}^2. \tag{3.1}$$

Proof. As $U(z)$ is regular in $\bar{\mathfrak{F}}^2$, it is also regular in an open convex domain $\mathfrak{F}_1^2 \supset \bar{\mathfrak{F}}^2$. By §2 above, the function $f(u)$ in (2.11) is a regular function of a complex variable u in $\mathfrak{F}_1^2$. By Runge's theorem, every regular function may be approximated by a polynomial. For every $\epsilon > 0$, there exists, therefore, a polynomial

$$p_n(u) = a_0 + \sum_{k=1}^{n} (a_{2k-1} - ia_{2k})u^k$$

such that

$$|f(u) - p_n(u)| \leqq \epsilon/4\pi C_1, \qquad u \in \bar{\mathfrak{F}}^2, \qquad C_1 = \max_{\substack{-1 \leq t \leq 1 \\ z \in \bar{\mathfrak{F}}^2}} |\mathbf{E}(z; t)|. \tag{3.2}$$

By (2.11) and (2.12)

$$\begin{aligned} &\left| U(z) - \sum_{k=0}^{2n} a_k P_k(z) \right| \\ &\leqq 2 \int_{-1}^{1} |\mathbf{E}(z; t)| \cdot \left| f\left(z \cdot \frac{1 - t^2}{2}\right) - p_n\left(z \cdot \frac{1 - t^2}{2}\right) \right| (1 - t^2)^{-\frac{1}{2}} dt \\ &\leqq \frac{\epsilon}{2\pi} \int_{-1}^{1} (1 - t^2)^{-\frac{1}{2}} dt = \epsilon. \end{aligned} \tag{3.3}$$

THEOREM 3. *Let* $\mathfrak{F}^2$ *be a domain in which there exists no solution of* (2.10) *which vanishes on the boundary other than* $U \equiv 0$. *Let* $U(z)$ *satisfy the equation* (2.10) *in* $\mathfrak{F}^2$ *and on* $\mathfrak{f}^1$ *possess boundary values* $\phi(s)$ *which have continuous second derivatives. Then for every* $\epsilon > 0$, *there exists a linear combination* W *such that* (3.1) *is valid in the* (*closed*) *domain* $\bar{\mathfrak{F}}^2$.

Theorem 3 follows from Lemma 1 and

LEMMA 2. *Every solution* $U(z)$ *of* (2.10) *with the properties indicated in Theorem 3 can be approximated in* $\bar{\mathfrak{F}}^2$ *by another solution* $V(z)$ *of* (2.10) *which is regular in* $\bar{\mathfrak{F}}^2$.

Proof. I. By the continuity of U in $\bar{\mathfrak{F}}^2$, for every $\epsilon > 0$, there exists a δ_0 such that for $0 \leqq \delta \leqq \delta_0$

$$|U(z) - R(z)| \leqq \tfrac{1}{2}\epsilon, \qquad z \in \bar{\mathfrak{F}}^2, \qquad R(z) = U\left(\frac{z}{1 + \delta}\right). \tag{3.4}$$

$R(z)$ satisfies the equation

$$\begin{aligned} &R_{z\bar{z}} + 2 \operatorname{Re}\left[(1 + \delta)^{-1} A\left(\frac{z}{1 + \delta}\right) R_z\right] \\ &\qquad + (1 + \delta)^{-2} C\left(\frac{z}{1 + \delta}\right) R \equiv \mathbf{L}(R) + F_\delta(z) = 0 \end{aligned} \tag{3.5}$$

in $\mathfrak{F}_\delta^2 = (1 + \delta)\mathfrak{F}^2$, where[7]

$$F_\delta(z) = 2 \operatorname{Re} \left\{ \left[(1 + \delta)^{-1} A \left(\frac{z}{1 + \delta} \right) - A(z) \right] R_z \right\} \tag{3.6}$$
$$+ \left[(1 + \delta)^{-2} C \left(\frac{z}{1 + \delta} \right) - C(z) \right] R.$$

By our hypothesis that there exists no "fundamental solution" of (2.10) in $\mathfrak{F}^2$, the functions U_z and $U_{\bar{z}}$ are continuous in $\bar{\mathfrak{F}}^2$ by Lichtenstein's theorem ([10], p. 1285).

Since $\tilde{A} \equiv A\left(\frac{z}{1 + \delta}\right)$ and $\tilde{B}$, $\tilde{C}$, R, and their derivatives are bounded in $\bar{\mathfrak{F}}^2$, we can, moreover, choose δ_0 so small that

$$| F_\delta(z) | \leqq \epsilon / 2C_2 , \qquad z \in \bar{\mathfrak{F}}^2, \tag{3.7}$$

where C_2 is a constant[8] to be determined later.

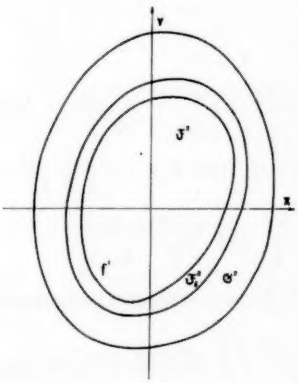

FIG. 1

II. In order to complete the proof of Lemma 2 we require

LEMMA 3. *Let $S_\delta(z)$ be a function which in $\mathfrak{F}_\delta^2$ satisfies the equation*

$$\mathbf{L}(S_\delta) - F_\delta = 0, \tag{3.8}$$

and vanishes on the boundary $\mathfrak{f}_\delta^1$ of $\mathfrak{F}_\delta^2$. Moreover, suppose

$$| F_\delta(z) | \leqq M, \qquad z \in \bar{\mathfrak{F}}_\delta^2 . \tag{3.9}$$

Then there exists a constant C_2 depending on $\mathbf{L}$ *and $\mathfrak{F}^2$ (but neither on S_δ nor on δ) such that*

$$| S_\delta(z) | \leqq C_2 M. \tag{3.10}$$

[7] We designate by $(1 + \delta)\mathfrak{F}^2$ the domain obtained from $\mathfrak{F}^2$ by the transformation $z^* = (1 + \delta)z$.

[8] The C_k ($k = 1, 2, \cdots, 21$) are appropriate constants.

Proof. A solution $S_\delta(z)$ of the elliptic equation satisfies the integral equation

$$S_\delta(z) = \frac{1}{2\pi} \iint_{\mathfrak{F}_\delta^2} K_\delta(z, \zeta) S_\delta(\zeta)\, d\omega_\zeta + N_\delta(z), \tag{3.11}$$

$$z \in \mathfrak{F}_\delta^2,\ d\omega = d\xi\, d\eta,\ \zeta = \xi + i\eta,$$

$$N_\delta(z) = \frac{2}{\pi} \iint_{\mathfrak{F}_\delta^2} F_\delta(\zeta) G\left(\frac{z}{1+\delta}, \frac{\zeta}{1+\delta}\right) d\omega_\zeta, \tag{3.12}$$

where

$$K_\delta(z, \zeta) = 4C(\zeta) G\left(\frac{z}{1+\delta}, \frac{\zeta}{1+\delta}\right) - 8 \operatorname{Re}\left\{\left[A(\zeta) G\left(\frac{z}{1+\delta}, \frac{\zeta}{1+\delta}\right)\right]_\zeta\right\} \tag{3.13}$$

and $G(z, \zeta)$ is the Green's function of $\mathfrak{F}^2$ (see [8], p. 536).

In order to avoid the inconvenience of operating with varying domains $\mathfrak{F}_\delta^2$ we introduce, instead of z, ζ,

$$Z = \frac{z}{1+\delta}, \qquad \mathrm{Z} = \frac{\zeta}{1+\delta}, \tag{3.14}$$

$$S_\delta^*(Z) = S_\delta[Z(1+\delta)], \qquad N_\delta^*(\mathrm{Z}) = N_\delta[\mathrm{Z}(1+\delta)], \tag{3.15}$$

$$\begin{aligned} K_\delta^*(Z, \mathrm{Z}) &= 4C(\mathrm{Z}(1+\delta)) G(Z, \mathrm{Z}) - \frac{8}{1+\delta} \operatorname{Re}\{[A(\mathrm{Z}(1+\delta)) G(Z, \mathrm{Z})]_{\mathrm{Z}}\} \\ &= \frac{\Lambda(Z, \mathrm{Z})}{|Z - \mathrm{Z}|} + \delta \cdot \frac{\mathrm{P}_1(Z, \mathrm{Z}; \delta)}{|Z - \mathrm{Z}|}. \end{aligned} \tag{3.16}$$

It is well known that $S_\delta^*(Z)$ also satisfies the equation

$$S_\delta^*(Z) = \frac{(1+\delta)^6}{8\pi^3} \iint_{\mathfrak{F}_\delta^2} K_\delta^{*(3)}(Z, \mathrm{Z}) S_\delta^*(\mathrm{Z})\, d\omega_{\mathrm{Z}} + N_\delta^{*(3)}(Z), \qquad Z \in \mathfrak{F}^2, \tag{3.17}$$

where

$$\begin{aligned} N_\delta^{*(3)}(Z) &= N_\delta^*(Z) + \frac{(1+\delta)^2}{2\pi} \iint_{\mathfrak{F}_\delta^2} K_\delta^*(Z, \mathrm{Z}) N_\delta^*(\mathrm{Z})\, d\omega_{\mathrm{Z}} \\ &\quad + \frac{(1+\delta)^4}{4\pi^2} \iint_{\mathfrak{F}_\delta^2} K_\delta^{*(2)}(Z, \mathrm{Z}) N_\delta^*(\mathrm{Z})\, d\omega_{\mathrm{Z}} \end{aligned} \tag{3.18}$$

(see [8], p. 356).

We shall show that the expression $D_\delta^{(3)}[\frac{1}{8}(1+\delta)^6 \pi^{-3}]$ which is contained in the formula for the resolvent of (3.17)

$$\Gamma_\delta^{(3)}\left(Z, \mathrm{Z}; \frac{(1+\delta)^6}{8\pi^3}\right)$$

(see [8], p. 373) depends continuously on δ. For this reason, we consider the iterated functions $K_\delta^{*(n)}(Z, \mathrm{Z})$ $(n = 2, 3)$. Since A_δ and C are continuous and therefore bounded, and since

$$|Z - \mathrm{Z}|G(Z, \mathrm{Z}) = |Z - \mathrm{Z}|\,[\log|Z - \mathrm{Z}| + T(Z, \mathrm{Z})]$$

and

$$|Z - \mathrm{Z}|G_{\mathrm{Z}}(Z, \mathrm{Z}) = |Z - \mathrm{Z}|\left[\frac{\partial \log|Z - \mathrm{Z}|}{\partial \mathrm{Z}} + \frac{\partial T(Z, \mathrm{Z})}{\partial \mathrm{Z}}\right]$$

are likewise bounded[9] in $\mathfrak{F}^2$, there exists a constant C_4 such that (see (3.16))

$$|\Lambda(Z, \mathrm{Z})| \leqq C_4, \qquad |\mathrm{P}_1(Z, \mathrm{Z}; \delta)| \leqq C_4, \qquad Z \in \mathfrak{F}^2, \mathrm{Z} \in \mathfrak{F}^2. \tag{3.19}$$

Therefore,

$$\begin{aligned} K_\delta^{*(2)}(Z, \mathrm{Z}) &= K_0^{*(2)}(Z, \mathrm{Z}) \\ &\quad + \delta \iint_{\mathfrak{F}_\delta^2} \frac{2\Lambda(Z, S)\mathrm{P}_1(S, \mathrm{Z}; \delta) + \delta \mathrm{P}_1(Z, S; \delta)\mathrm{P}_1(S, \mathrm{Z}; \delta)}{|Z - S|\cdot|S - \mathrm{Z}|}\, d\omega_S. \end{aligned} \tag{3.20}$$

Since $|\Lambda| \leqq C_4$ and $|\mathrm{P}_1| \leqq C_4$, we obtain by a standard procedure ([8], p. 362)

$$K_\delta^{*(2)}(Z, \mathrm{Z}) = K_0^{*(2)}(Z, \mathrm{Z}) + \delta \mathrm{P}_2(Z, \mathrm{Z}; \delta) \log|Z - \mathrm{Z}|, \tag{3.21}$$

where

$$|\mathrm{P}_2(Z, \mathrm{Z}; \delta)| \leqq C_5, \qquad Z \in \mathfrak{F}^2, \mathrm{Z} \in \mathfrak{F}^2. \tag{3.22}$$

[9] The fact that this expression is bounded can be proved as follows: Let $w(z)$ be the function which transforms $\mathfrak{F}^2$ into $|w| < 1$, taking the origin and the real axis into themselves. We write

$$\begin{aligned} Q(z, \zeta) = \tfrac{1}{2}\{&\log[w(z) - w(\zeta)] - \log[1 - \overline{w(z)}w(\zeta)] \\ &+ \log[\overline{w(z)} - \overline{w(\zeta)}] - \log[1 - w(z)\overline{w(\zeta)}]\}. \end{aligned}$$

We then obtain

$$\begin{aligned} \left|\frac{dQ(z, \zeta)}{d\zeta}(z - \zeta)\right| &= \left|\frac{dQ(z, \zeta)}{dw(\zeta)}\right|\cdot|w(z) - w(\zeta)|\cdot|w'(\zeta)|\cdot\left|\frac{z - \zeta}{w(z) - w(\zeta)}\right| \\ &= \left|\left[\frac{-1}{w(z) - w(\zeta)} + \frac{\overline{w(z)}}{1 - \overline{w(z)}w(\zeta)}\right]\right|\cdot\left|\frac{z - \zeta}{w(z) - w(\zeta)}\right|\cdot|w(z) - w(\zeta)|\cdot|w'(\zeta)| \\ &\leqq \left[-1 + \left|\overline{w(z)}\left(\frac{w(z) - w(\zeta)}{1 - \overline{w(z)}w(\zeta)}\right)\right|\right]\cdot|w'(\zeta)|\cdot\left|\frac{z - \zeta}{w(z) - w(\zeta)}\right|. \end{aligned}$$

Since $(w(z) - w(\zeta))/(1 - \overline{w(z)}w(\zeta))$ transforms the circle $|w| < 1$ into itself, we find that the absolute value of this fraction is <1, and according to a theorem due to Kellogg, $|w'(\zeta)|$ and $|(z - \zeta)/(w(z) - w(\zeta))|$ are bounded [12].

By repeating these same processes, we obtain

$$(3.23)\qquad \begin{aligned} K_\delta^{*(3)}(Z, \mathrm{Z}) &= K_0^{*(3)}(Z, \mathrm{Z}) + \delta \mathrm{P}_3(Z, \mathrm{Z}; \delta); \\ |\, \mathrm{P}_3(Z, \mathrm{Z}; \delta) \,| &\leqq C_6, \qquad Z \in \bar{\mathfrak{F}}^2, \mathrm{Z} \in \bar{\mathfrak{F}}^2. \end{aligned}$$

$K_0^{*(3)}(Z, \mathrm{Z})$ being finite, there exists a constant C_7 such that for $0 \leqq \delta \leqq \delta_0$ (δ_0 sufficiently small)

$$(3.24)\qquad |\, K_\delta^{*(3)}(Z, \mathrm{Z}) \,| \leqq C_7 .$$

The series for $D_\delta^{(3)}[\frac{1}{8}(1 + \delta)^6\pi^{-3}]$ (see [8], p. 371, (5)) converges uniformly in δ and from the uniform continuity of $K_\delta^{*(3)}(Z, \mathrm{Z})$ follows the continuity of $D_\delta^{(3)}[\frac{1}{8}(1 + \delta)^6\pi^{-3}]$ considered as a function of δ. Since $D_0^{(3)}(\frac{1}{8}\pi^{-3}) = a \neq 0$, there exists an interval $\{0 \leqq \delta \leqq \delta_1\}$ such that, for $\delta \leqq \delta_1 \leqq \delta_0$, $|\, D_\delta^{(3)}[\frac{1}{8}(1 + \delta)^6\pi^{-3}] \,| \geqq \frac{1}{2} \,|\, a \,|$. It follows from (3.24) that

$$D_\delta^{(3)}\left[\begin{matrix} Z \\ \mathrm{Z} \end{matrix}\,\middle|\, \frac{(1 + \delta)^6}{8\pi^3}\right]$$

is a uniformly bounded function ([8], p. 372, (9)); hence we obtain for the resolvent of (3.17)

$$(3.25)\qquad \left|\, \Gamma_\delta^{(3)}\left(Z, \mathrm{Z}; \frac{(1 + \delta)^6}{8\pi^3}\right) \right| \leqq C_8, \qquad Z \in \bar{\mathfrak{F}}^2, \mathrm{Z} \in \bar{\mathfrak{F}}^2.$$

Since $K_\delta^*(Z, \mathrm{Z})$ and $K_\delta^{*(2)}(Z, \mathrm{Z})$ tend to infinity like $1/|\, Z - \mathrm{Z} \,|$ and $\log |\, Z - \mathrm{Z} \,|$ respectively, it follows from (3.9), (3.12), and (3.18) that there exists a constant C_9 such that

$$(3.26)\qquad |\, N_\delta^*(Z) \,| \leqq C_9 M, \qquad Z \in \bar{\mathfrak{F}}^2.$$

On the other hand ([8], p. 373, (13)) we have

$$(3.27)\qquad \begin{aligned} |\, S_\delta^*(Z) \,| &= |\, N_\delta^*(Z) \,| + \frac{(1 + \delta)^6}{8\pi^3} \iint_{\mathfrak{F}^2} \Gamma_\delta^{(3)}\left(Z, \mathrm{Z}; \frac{(1 + \delta)^6}{8\pi^3}\right) N^*(\mathrm{Z})\, d\omega_{\mathrm{Z}} \\ &\leqq [C_9 + \tfrac{1}{8}(1 + \delta)^6\pi^{-3} C_8 C_9 \mathrm{A}(\mathfrak{F}^2)] M, \end{aligned}$$

$\mathrm{A}(\mathfrak{F}^2)$ being the area of $\mathfrak{F}^2$. By setting C_2 equal to this coefficient of M, with $\delta = \delta_1$, we obtain (3.10).

III. By (3.5) and (3.8) we see that $(R + S_\delta)$ satisfies the equation $\mathbf{L}(R + S_\delta) = 0$ in $\mathfrak{F}_\delta^2$ and by a theorem of Picard ([10], p. 1320) it is thus regular in $\bar{\mathfrak{F}}^2 \subset \mathfrak{F}_\delta^2$. Moreover, by (3.4), (3.7), and (3.10) we find that

$$(3.28)\qquad |\, U - (R + S_\delta) \,| \leqq |\, U - R \,| + |\, S_\delta \,| \leqq \epsilon.$$

By choosing ϵ arbitrarily small, we obtain Lemma 2.

4. **First boundary problem.** In this section we suppose that the equation (2.10) has no fundamental solutions in $\mathfrak{F}^2$; i.e., that $U \equiv 0$ is the only solution of (2.10) which vanishes on the boundary $\mathfrak{f}^1$.

THEOREM 4. *Let $\phi(s)$ be the boundary values of a solution $U(z)$ of (2.10) and suppose that $d^2\phi(s)/ds^2$ is continuous. Let the coefficients $A_\nu^{(n)}$ in the expression*

$$U_n(z) = \sum_{\nu=0}^{n} A_\nu^{(n)} P_\nu(z) \tag{4.1}$$

be determined in such a manner that for every n the quantity

$$\max_{z \in \mathfrak{f}^1} | U(z) - U_n(z) | \tag{4.2}$$

will be a minimum. Then

$$U(z) = \lim_{n \to \infty} U_n(z), \qquad z \in \bar{\mathfrak{F}}^2. \tag{4.3}$$

Proof. I. We denote the minimum of (4.2) by J_n. By Theorem 3 it follows that

$$\lim_{n \to \infty} J_n = 0, \tag{4.4}$$

since for every $\epsilon > 0$ we can find an expression $\sum_{\nu=1}^{n} B_\nu P_\nu(z)$ such that in $\bar{\mathfrak{F}}^2$ (and therefore on $\mathfrak{f}^1$) $| U(z) - \sum_{\nu=1}^{n} B_\nu P_\nu(z) | \leqq \epsilon$. Since $J_n \leqq \epsilon \mathrm{L}(\mathfrak{f}^1)$, $\mathrm{L}(\mathfrak{f}^1)$ being the length of $\mathfrak{f}^1$ and ϵ an arbitrary positive number, (4.4) follows.

II. $V_n(z) = U(z) - U_n(z)$ satisfies the equation (2.10) in $\mathfrak{F}^2$; therefore

$$V_n(z) = \frac{1}{2\pi} \iint_{\mathfrak{F}^2} K_0(z, \zeta) V_n(\zeta)\, d\omega_\zeta + N(z), \tag{4.5}$$

where

$$N(z) = \frac{1}{2\pi} \iint_{\mathfrak{F}^2} V_n(\zeta) \frac{\partial G(z, \zeta)}{\partial \mathrm{n}_\zeta}\, ds_\zeta$$

(n_ζ being the interior normal) is the harmonic function which takes the value[8] $V_n(z)$ on $\mathfrak{f}^1$ and K_0 is given by (3.13) with $\delta = 0$. Since

$$| V_n(z) | \leqq \epsilon, \qquad z \in \mathfrak{f}^1, \tag{4.6}$$

it follows that

$$| N(z) | \leqq \epsilon, \qquad z \in \bar{\mathfrak{F}}^2. \tag{4.7}$$

A procedure analogous to that of §3 leads us to the result that in $\bar{\mathfrak{F}}^2$

$$| U(z) - U_n(z) | = | V_n(z) | \leqq \epsilon C_{10}. \tag{4.8}$$

Since ϵ can be chosen arbitrarily small and C_{10} is independent of n, we obtain (4.3).

Remark. In some cases we may replace (4.2) by

$$\int_{\mathfrak{f}^1} | U(z) - U_n^*(z) |^p\, ds, \qquad U_n^*(z) = \sum_{\nu=0}^{n} B_\nu^{(n)} P_\nu(z), \qquad p > 1. \tag{4.9}$$

(The numerical determination of the $B_\nu^{(n)}$ is easier than that of the $A_\nu^{(n)}$. We can use (4.9) when the Green's function $H(z, \zeta)$ of equation (2.10) for the domain $\mathfrak{F}^2$ possesses the property that

$$\int_{\mathfrak{f}^1} \left| \frac{\partial H(z, \zeta)}{\partial \mathfrak{n}_\zeta} \right|^{p'} ds \leqq C_{11}(z) < \infty, \qquad \frac{1}{p} + \frac{1}{p'} = 1, \qquad z \in \mathfrak{F}^2. \tag{4.10}$$

In this case,

$$U(z) = \lim_{n\to\infty} U_n^*(z), \qquad z \in \mathfrak{F}^2. \tag{4.11}$$

Proof. We denote the minimum of (4.9) by I_n; since $I_n \leqq J_n^p L(\mathfrak{f}^1)$, it follows that $\lim_{n\to\infty} I_n = 0$, and using

$$\begin{aligned} |U(z) - U_n^*(z)| &\leqq \left[\int_{\mathfrak{f}^1} \left| \frac{\partial H(z, \zeta)}{\partial \mathfrak{n}_\zeta} \right|^{p'} ds_\zeta \right]^{1/p'} \cdot \left[\int_{\mathfrak{f}^1} |U - U_n^*|^p ds_\zeta \right]^{1/p} \\ &\leqq (I_n)^{1/p}(C_{11}(z))^{1/p'}, \end{aligned} \tag{4.12}$$

we obtain (4.11).

Remark. In the case of $\mathfrak{F}^2$ considered in this paper, hypothesis (4.10) is fulfilled for every $p' > 1$, since by [9] there exists for every operator $\mathbf{L}$ and every $\mathfrak{F}^2$ a function $N(z, \zeta)$ which at the point $\zeta = z$ has the same behavior as $\log |z - \zeta|$ and which is regular in a domain $\mathfrak{G}^2 \supset \bar{\mathfrak{F}}^2$ except for this point. Therefore the second derivatives of $N(z, \zeta)$ are continuous on $\mathfrak{f}^1$. By Lichtenstein's theorem, the solution $V(\zeta)$ of the adjoint equation of $\mathbf{L}$ which assumes the values $N(z, \zeta)$ on $\mathfrak{f}^1$ possesses continuous derivatives in $\bar{\mathfrak{F}}^2$. Since $H(z, \zeta) = N(z, \zeta) + V(\zeta)$, our assumption follows.

5. **Determination of the characteristic values.** A parameter λ now appears in the equation and values of λ (characteristic values[10]) are sought for which there exist solutions of the equation which vanish on $\mathfrak{f}^1$ without vanishing identically.

We shall consider equations of the form

$$\mathbf{L}_1(U) \equiv U_{z\bar{z}} + 2 \operatorname{Re} [AU_z] + [C + \lambda E]U = 0, \tag{5.1}$$

where $E = E(z)$ is positive and regular in a sufficiently large domain. Since $\mathrm{E}^*(z; t)$ depends continuously on λ (see [6], pp. 1173–1176), the functions $P_n(z) = P_n(z; \lambda)$ defined by (2.12) also depend continuously on λ. It is well known that (under appropriate hypotheses) there exists a discrete set of characteristic values λ_m $(m = 1, 2, \cdots)$ and corresponding fundamental solutions $\Phi_m(z)$ of (5.1) with $\iint_{\mathfrak{F}^2} E\Phi_m^2 \, d\omega = 1$. In order to determine λ_m and $\Phi_m(z)$, we calculate $A_\nu^{(n)}$ and $\tau^{(n)}$ so that

[10] The characteristic values of the equation (5.1) differ from those of the associated integral equation (6.4) by the factor $(2\pi)^{-1}$ (see [8], pp. 535, 536). By the characteristic values we understand here and in the following those of the differential equation.

$$(5.2) \qquad \max_{z \in \mathfrak{l}^1} \left| \sum_{\nu=0}^{n} A_\nu^{(n)} P_\nu(z; \tau^{(n)}) \right|$$

will be a minimum, subject to the condition that

$$(5.3) \qquad \mathrm{N}\left[\sum_{\nu=0}^{n} A_\nu^{(n)} P_\nu(z; \tau^{(n)})\right] \equiv \iint_{\mathfrak{F}^2} \left[\sum_{\nu=0}^{n} A_\nu^{(n)} P_\nu(z; \tau^{(n)})\right]^2 E(z)\, d\omega = 1.$$

We shall designate this minimum by $\mathbf{P}_n$. For the determination of $\tau^{(n)}$ and $A_\nu^{(n)}$ we obtain a transcendental equation which possesses several solutions.

THEOREM 5. *If, in using the procedure indicated above, we choose*[11] $\tau^{(n)}$ *in an interval* $\mathfrak{t}^1 = [\tau^* \leqq \tau \leqq \tau^{**}]$ *which contains one and only one characteristic value* λ_m *of* (5.1) *for each domain* $\mathfrak{F}^2$, *then*

$$(5.4) \qquad \lim_{n \to \infty} \tau^{(n)} = \lambda_m.$$

In order to prove this, we consider the expression

$$(5.5) \qquad M(\rho) = \lim_{n \to \infty} M_n(\rho),$$

where $M_n(\rho)$ denotes the minimum of

$$(5.6) \qquad \max_{z \in \mathfrak{l}^1} \left| \sum_{\nu=0}^{n} B_\nu^{(n)} P_\nu(z; \rho) \right|$$

for a fixed ρ, subject to the condition that $\mathrm{N}\left[\sum_{\nu=0}^{n} B_\nu^{(n)} P_\nu(z; \rho)\right] = 1$.

LEMMA 4. *For every characteristic value* λ_m *we have*

$$(5.7) \qquad \lim_{\rho \to \lambda_m} M(\rho) = 0.$$

Proof. Suppose that $\mathfrak{G}^2 \supset \bar{\mathfrak{F}}^2$ (see Figure 1) and that the characteristic values of (5.1) corresponding to $\mathfrak{G}^2$ do not belong to a closed interval $[\lambda_m, \lambda^*]$. We write

$$(5.8) \qquad \mathbf{T}(U) \equiv U_{z\bar{z}} + 2\,\mathrm{Re}\,[AU_z] + [C + \rho^{(n)}E]U = 0, \qquad \rho^{(n)} \in [\lambda_m, \lambda^*].$$

By Lemma 3 there exists a constant[12] C_2 such that for every solution $P(z)$ of

$$(5.9) \qquad \mathbf{T}[P(z)] + G(z) = 0$$

vanishing on the boundary of $\mathfrak{G}^2$, we have

$$(5.10) \qquad |P(z)| \leqq C_2 \max_{\zeta \in \bar{\mathfrak{G}}^2} |G(\zeta)|.$$

[11] As we shall show, this is always possible for n sufficiently large.

[12] Since $C_2(\tau)$, $\tau \in \{\lambda_m, \lambda^*\}$ is continuous and positive, there exists a $C_2^* > 0$ such that $C_2^* \leqq C_2(\tau)$, $\tau \in \{\lambda_m, \lambda^*\}$.

For every $\epsilon > 0$, we can choose a δ_0 so small and a $\rho^{(n)} \epsilon [\lambda_m, \lambda^*]$ so near to λ_m, $\rho^{(n)} \neq \lambda_m$, that

(i) for $\delta \leqq \delta_0$

$$|\Phi_m(z) - R(z)| \leqq \tfrac{1}{4}\epsilon, \quad z \in \bar{\mathfrak{F}}^2, \quad R(z) = \Phi_m\left(\frac{z}{1+\delta}\right); \tag{5.11}$$

(ii) $R(z)$ satisfies the equation

$$\mathbf{T}[R(z)] - F(z) = 0, \quad z \in \mathfrak{F}_\delta^2, \quad \mathfrak{F}_\delta^2 = (1+\delta)\mathfrak{F}^2, \tag{5.12}$$

and

$$|F(z)| \leqq \tfrac{1}{4}\epsilon C_2, \quad C_2 = C_2(\mathbf{T}, \mathfrak{G}^2), \tag{5.13}$$

where

$$\begin{aligned} F(z) = 2\,\mathrm{Re}\left\{\left[(1+\delta)^{-1}A\left(\frac{z}{1+\delta}\right) - A(z)\right]R_z\right\} \\ + \left[(1+\delta)^{-2}C\left(\frac{z}{1+\delta}\right) - C(z)\right]R + \left[\rho^{(n)}(1+\delta)^{-2}E\left(\frac{z}{1+\delta}\right) - \lambda_m E(z)\right]R. \end{aligned} \tag{5.14}$$

$F(z)$ is defined in $\mathfrak{F}^2$, but it is possible for $F(z)$ to be prolonged in $\mathfrak{G}^2$ so that it is three times differentiable there and so that (5.13) holds there.

We construct the function $S(z)$ satisfying the equation

$$\mathbf{T}[S(z)] + F(z) = 0 \tag{5.15}$$

in $\mathfrak{G}^2$ and vanishing on its boundary. By (5.10) and (5.13) we have in $\mathfrak{G}^2$ and so in $\bar{\mathfrak{F}}^2$,

$$|S(z)| \leqq \tfrac{1}{4}\epsilon. \tag{5.16}$$

$(R + S)$ satisfies the equation

$$\mathbf{T}[R(z) + S(z)] = 0, \quad z \in \mathfrak{F}^2. \tag{5.17}$$

Since the coefficients of $\mathbf{T}$ are regular functions of z, $\bar{z}$ in $\mathfrak{F}_\delta^2$, $(R + S)$ is also regular in $\mathfrak{F}_\delta^2$ ([10], p. 1320) and so in $\bar{\mathfrak{F}}^2$. On the other hand, by (5.11) and (5.16), we have

$$|\Phi_m - R - S| \leqq |\Phi_m - R| + |S| \leqq \tfrac{1}{2}\epsilon, \quad z \in \bar{\mathfrak{F}}^2. \tag{5.18}$$

Since $(R + S)$ is regular in $\bar{\mathfrak{F}}^2$, by Lemma 1 we can approximate to $(R + S)$ in $\bar{\mathfrak{F}}^2$ by an expression $\bar{P}_q = \sum_{\nu=0}^{q} B_\nu^{(q)} P_\nu(z; \rho^{(n)})$ in such a way that we have $|R + S - \bar{P}_q| \leqq \tfrac{1}{2}\epsilon$, $z \in \bar{\mathfrak{F}}^2$. Therefore, $|\Phi_m(z) - \bar{P}_q(z)| \leqq \epsilon$, $z \in \bar{\mathfrak{F}}^2$. Moreover, since $\iint_{\mathfrak{F}^2} E(z)\Phi_m^2(z)\, d\omega = 1$, we have

$$\begin{gathered} \iint_{\mathfrak{F}^2} E(z)\bar{P}_q^2(z)\, d\omega = 1 + \epsilon_2, \\ |\epsilon_2| \leqq C_{12}\,\epsilon, \quad C_{12} \leqq 2\iint_{\mathfrak{F}^2} |E(z)\Phi_m(z)|\, d\omega + \epsilon \iint_{\mathfrak{F}^2} E(z)\, d\omega, \end{gathered} \tag{5.19}$$

where $d\omega = dx\,dy$, and also

$$| \bar{P}_q(z)(1 + \epsilon_2)^{-\frac{1}{2}} | \leqq \epsilon(1 - C_{12}\epsilon)^{-1}, \qquad z \in \mathfrak{f}^1.$$

Since we can get an ϵ arbitrarily small by choosing $\rho^{(n)}$ sufficiently near to λ_m, (5.7) follows from $M(\rho^{(n)}) \leqq \epsilon/(1 - C_{12}\epsilon)^{\frac{1}{2}}$.

LEMMA 5. *Let $\mathfrak{p}^1$ be a closed interval containing no characteristic values λ_m $(m = 1, 2, \cdots)$. Then there exists a constant $C_{13} > 0$ such that for $\rho \in \mathfrak{p}^1$, we have*

$$M(\rho) \geqq C_{13}. \tag{5.20}$$

Proof. Suppose (5.20) is false. Then there is a sequence $\rho^{(p)}$ $(p = 1, 2, \cdots)$, $\lim_{p\to\infty} \rho^{(p)} = \tau_0 \in \mathfrak{p}^1$, and of particular solutions $V_\mu(z; \rho^{(p)}) = \sum_{\nu=1}^{\mu} a_\nu^{(\mu)} P_\nu(z, \rho^{(p)})$ satisfying (5.3) for which

$$\lim_{p\to\infty} \lim_{\mu\to\infty} \max_{z \in \mathfrak{f}^1} V_\mu(z; \rho^{(p)}) = 0. \tag{5.21}$$

Every V_μ also satisfies the integral equation

$$V_\mu(z; \rho^{(p)}) = \frac{1}{2\pi} \iint_{\mathfrak{F}^2} K(z, \zeta; \rho^{(p)})\, V_\mu(\zeta; \rho^{(p)})\, d\omega_\zeta + G_{\mu p}(z), \tag{5.22}$$

where

$$K(z, \zeta; \rho^{(p)}) = 4[C(\zeta) + \rho^{(p)} E(\zeta)]\, G(z, \zeta) - 8 \operatorname{Re} \{[A(\zeta) G(z, \zeta)]_\zeta\}$$

and

$$G_{\mu p}(z) = \frac{1}{2\pi} \int_{\mathfrak{f}^1} V_\mu(z; \rho^{(p)}) \frac{\partial G(z, \zeta)}{\partial n_\zeta}\, ds_\zeta$$

is the harmonic function which takes values $V_\mu(z; \rho^{(p)})$ on the boundary $\mathfrak{f}^1$ ([10], p. 1282). Since a harmonic function assumes its maximum on the boundary, it follows from (5.21) that

$$\lim_{p\to\infty} \lim_{\mu\to\infty} G_{\mu p}(z) = 0. \tag{5.23}$$

V_μ also satisfies the equation

$$V_\mu(z; \rho^{(p)}) = \frac{1}{8\pi^3} \iint_{\mathfrak{F}^2} K^{(3)}(z, \zeta; \rho^{(p)}) V_\mu(\zeta; \rho^{(p)})\, d\omega_\zeta + G_{\mu p}^{(3)}(z), \tag{5.24}$$

where

$$\begin{aligned} G_{\mu p}^{(3)}(z) = G_{\mu p}(z) &+ \frac{1}{2\pi} \iint_{\mathfrak{F}^2} K(z, \zeta; \rho^{(p)}) G_{\mu p}(\zeta)\, d\omega_\zeta \\ &+ \frac{1}{4\pi^2} \iint K^{(2)}(z, \zeta; \rho^{(p)})\, G_{\mu p}(\zeta)\, d\omega_\zeta \end{aligned} \tag{5.25}$$

([8], p. 356). By a procedure analogous to that previously used, it may be proved that there exists a neighborhood $\mathfrak{n}^1$ of τ_0 such that for $\rho^{(p)} \in \mathfrak{n}^1$, we have

$$K^{(2)}(z, \zeta; \rho^{(p)}) \leqq \Lambda_0 \log |z - \zeta| + \Lambda_1, \quad K^{(3)}(z, \zeta; \rho^{(p)}) \leqq \Lambda_2, \quad z \in \bar{\mathfrak{F}}^2, \zeta \in \bar{\mathfrak{F}}^2,$$

with all three Λ's finite.

Therefore, it follows from (5.23) that

$$\lim_{p \to \infty} \lim_{\mu \to \infty} G^{(3)}_{\mu p}(z) = 0. \tag{5.26}$$

Moreover, by methods previously used, we obtain

$$\Gamma^{(3)}(z, \zeta; \rho^{(p)}) \leqq C_{14}, \quad \rho^{(p)} \in \mathfrak{n}^1, \quad z \in \bar{\mathfrak{F}}^2, \zeta \in \bar{\mathfrak{F}}^2.$$

Since

$$V_\mu(z; \rho^{(p)}) = \iint_{\mathfrak{F}^2} \Gamma^{(3)}\left(z, \zeta; \rho^{(p)}; \frac{1}{8\pi^3}\right) G^{(3)}_{\mu p}(\zeta)\, d\omega_\zeta + G^{(3)}_{\mu p}(z), \quad z \in \bar{\mathfrak{F}}^2, \tag{5.27}$$

and $\Gamma^{(3)}$ is bounded, it follows from (5.26) that

$$\lim_{p \to \infty} \lim_{\mu \to \infty} V_\mu(z; \rho^{(p)}) = 0, \quad z \in \bar{\mathfrak{F}}^2. \tag{5.28}$$

Hence

$$\lim_{p \to \infty} \lim_{\mu \to \infty} \iint_{\mathfrak{F}^2} E(z) V^2_\mu(z; \rho^{(p)})\, d\omega = 0. \tag{5.29}$$

This is a contradiction of (5.3), and so (5.20) is proved.

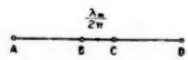

Fig. 2

Proof of Theorem 5. By Lemma 4 we have that

$$\lim_{n \to \infty} \mathrm{P}_n = 0, \tag{5.30}$$

where P_n is the function defined by (5.2). On the other hand,

$$\mathrm{P}_n = M_n(\tau^{(n)}) \tag{5.31}$$

(see (5.5)). By Lemma 5, to every closed interval $AB + CD$ which does not contain any characteristic value λ_m, there corresponds a positive δ such that for every n, we have

$$M_n(\tau) \geqq M(\tau) \geqq \delta, \quad \tau \in (AB + CD). \tag{5.32}$$

From (5.30), (5.31), and (5.32) it follows that for n sufficiently large, $\tau^{(n)}$ must belong to the complementary interval BC. Since an arbitrarily small neighborhood of λ_m can be chosen for BC, we obtain (5.4).

Remark. It is also possible to use, instead of (5.2), the minimum of

$$\int_{\mathfrak{l}^1} | W_n(z; \tau) |^k \, ds, \qquad W_n(z; \tau) = \sum_{\nu=0}^{n} B_\nu^{(n)} P_\nu(z; \tau), \qquad k \geqq 2, \tag{5.33}$$

subject to the condition that $\iint_{\mathfrak{F}^2} E(z) W_n^2(z) \, d\omega = 1$. We denote this minimum by R_n. Let $M^{(1)}(\rho) = \lim_{n \to \infty} M_n^{(1)}(\rho)$, where $M_n^{(1)}(\rho)$ is the minimum of (5.33), for W_n subject to the condition given and for fixed ρ. Since $M_n^{(1)}(\rho) \leqq [M_n(\rho)]^k L(\mathfrak{l}^1)$, $L(\mathfrak{l}^1)$, as before, being the length of $\mathfrak{l}^1$, it follows that $\lim_{\rho \to \lambda_m} M^{(1)}(\rho) = 0$ for every value of λ_m. On the other hand, it follows from the relation

$$\begin{aligned} \Gamma(z, \zeta; \bar{\rho}) &= K(z, \zeta) + \bar{\rho} K^{(2)}(z, \zeta) + \bar{\rho}^2 \left[D_3\left(\begin{matrix} z \\ \zeta \end{matrix} \,\middle|\, \bar{\rho}^3\right) \Big/ D_3(\bar{\rho}^3) \right] \\ &\quad + \bar{\rho}^3 \iint_{\mathfrak{F}^2} [K(z, \zeta) + \bar{\rho} K^{(2)}(z, \zeta)] \left[D_3\left(\begin{matrix} z \\ \zeta \end{matrix} \,\middle|\, \bar{\rho}^3\right) \Big/ D_3(\bar{\rho}^3) \right] d\omega_\zeta, \end{aligned} \tag{5.34}$$

where $\bar{\rho} = \rho/2\pi$ (see [8], p. 382, (33)), that for every $\rho \neq \lambda_m$ we have $| z - \zeta | \, \Gamma(z, \zeta; \bar{\rho}) \leqq C(\rho) < \infty$. By methods analogous to those of Lemma 3, we can prove that to every interval $\mathfrak{n}^1$, there corresponds a constant C_{15} such that $C(\rho) \leqq C_{15}$, $\rho \in \mathfrak{n}^1$.

As above, we obtain

$$W_n(z; \tau^{(p)}) = \frac{1}{2\pi} \iint_{\mathfrak{F}^2} K(z, \zeta; \tau^{(p)}) W_n(\zeta; \tau^{(p)}) \, d\omega_\zeta + F_p(z), \tag{5.35}$$

where $F_p(z)$ is the harmonic function which takes the values $W_n(z; \tau^{(p)})$ on $\mathfrak{l}^1$. From (5.35) we obtain

$$W_n(z; \tau^{(p)}) = F_p(z) + \frac{1}{2\pi} \iint_{\mathfrak{F}^2} \Gamma\left(z, \zeta; \tau^{(p)}; \frac{1}{2\pi}\right) F_p(\zeta) \, d\omega_\zeta, \tag{5.36}$$

$$\begin{aligned} | W_n(z; \tau^{(p)}) | &\leqq | F_p(z) | + \frac{1}{2\pi} \left\{ \iint_{\mathfrak{F}^2} \left[\Gamma\left(z, \zeta; \tau^{(p)}; \frac{1}{2\pi}\right) \right]^{k'} d\omega_\zeta \right\}^{1/k'} \\ &\quad \cdot \left\{ \iint_{\mathfrak{F}^2} | F_p(\zeta) |^k \, d\omega_\zeta \right\}^{1/k}, \qquad k > 2, \end{aligned} \tag{5.37}$$

where $1/k + 1/k' = 1$, so that

$$| W_n(z; \tau^{(p)}) | \leqq \begin{cases} | F_p(z) | + C_{16} \, T(z) & \text{for } k = 2, \\ | F_p(z) | + C_{17} \, R_n^{1/k} & \text{for } k > 2, \end{cases} \tag{5.38}$$

where

$$T(z) = \left[\iint_{\mathfrak{F}^2} | z - \zeta |^{-2\epsilon} F_p^2(\zeta) \, d\omega_\zeta \right]^{\frac{1}{2}},$$

and

$$\iint_{\mathfrak{F}^2} E(z) W_n^2 \, d\omega \leqq [\max_{z \in \mathfrak{F}^2} | E(z) |] \Big\{ R_n + C_{16}^2 R_n \cdot \max_{\zeta \in \mathfrak{F}^2} \left[\iint_{\mathfrak{F}^2} | \zeta - z |^{-2\epsilon} d\omega \right] \tag{5.39}$$
$$+ 2 \left[C_{16} \iint_{\mathfrak{F}^2} | F_p(z) | [T^2(z) + 1] \, d\omega \right] \Big\}.$$

As has been proved in [5] (p. 104), for every domain $\mathfrak{F}^2$ there exists a constant C_{18} such that for every harmonic function, hence for $F_p(z)$, we have

$$\iint_{\mathfrak{F}^2} | F_p(z) |^k \, d\omega \leqq C_{18} \int_{\mathfrak{f}^1} | F_p(z) |^k \, ds = C_{18} \int_{\mathfrak{f}^1} | W_n(z; \tau^{(p)}) |^k \, ds = C_{18} R_n. \tag{5.40}$$

From (5.39) and (5.40) we obtain

$$\lim_{p \to \infty} \lim_{n \to \infty} \iint_{\mathfrak{F}^2} [W_n(z; \tau^{(p)})]^2 E(z) \, d\omega = 0, \tag{5.41}$$

and this contradicts $\iint_{\mathfrak{F}^2} W_n^2 E(z) \, d\omega = 1$.

6. Determination of fundamental solutions.

THEOREM 6. *If we suppose in addition to the hypothesis of Theorem 5 that*

(1) (5.1) *is a self-adjoint equation,*

(2) *to* λ_m *belongs only one fundamental solution* $\Phi_m(z)$ *with* $\iint_{\mathfrak{F}^2} E(z) \Phi_m^2(z) \, d\omega = 1$,

then we have

$$\Phi_m(z) = \lim_{n \to \infty} \sum_{\nu=0}^{n} A_\nu^{(n)} P_\nu(z; \tau^{(n)}), \qquad z \in \bar{\mathfrak{F}}^2. \tag{6.1}$$

Proof. Since (5.1) is a self-adjoint equation, it may be written in the form

$$U_{z\bar{z}} + CU + \lambda EU = 0, \qquad E > 0, \tag{6.2}$$

and there exists a function $H(z, \zeta)$ which satisfies the equation

$$U_{z\bar{z}} + CU = 0 \tag{6.3}$$

for $z \in \mathfrak{F}^2$, except at the point ζ, which vanishes on $\mathfrak{f}^1$, and which behaves like $\log | z - \zeta |$ in the neighborhood of $z = \zeta$. $H(z, \zeta)$, called the Green's function

of (6.3) for $\mathfrak{F}^2$, is symmetric in the pair of variables $(z, \bar{z})$ and $(\zeta, \bar{\zeta})$ ([10], p. 1289).

$$W_n(z; \tau^{(n)}) = \sum_{\nu=0}^{n} A_\nu^{(n)} P_\nu(z, \tau^{(n)})$$

(see (5.2)) satisfies the integral equation

$$W_n(z; \tau^{(n)}) = \frac{\tau^{(n)}}{2\pi} \iint_{\mathfrak{F}^2} H(z, \zeta)E(\zeta)W_n(\zeta; \tau^{(n)})\, d\omega_\zeta + F_n(z), \tag{6.4}$$

where

$$F_n(z) = \frac{1}{2\pi} \int_{\mathfrak{f}^1} W_n(\zeta; \tau^{(n)}) \frac{\partial H(z, \zeta)}{\partial \mathfrak{n}_\zeta}\, ds_\zeta \tag{6.5}$$

is the solution of (6.3) taking on $\mathfrak{f}^1$ the values of $W_n(z; \tau^{(n)})$ ([10], p. 1289, (7)). Since there is no solution of (6.3), other than $U \equiv 0$, which vanishes on $\mathfrak{f}^1$, we can prove in a way analogous to that of §3:

Lemma 6. *Let $L(z)$ satisfy (6.3) in $\mathfrak{F}^2$. There exists a constant C_{19}, depending only on (6.3) and $\mathfrak{F}^2$, so that*

$$|L(z)| \leqq C_{19} \max_{\zeta \in \mathfrak{f}^1} |L(\zeta)|, \qquad z \in \bar{\mathfrak{F}}^2. \tag{6.6}$$

It follows, therefore, from (5.30) that

$$\lim_{n\to\infty} F_n(z) = 0, \qquad z \in \bar{\mathfrak{F}}^2. \tag{6.7}$$

In the case of a kernel of the form $E(z)H(z, \zeta)$, one introduces another kernel $S_1(z, \zeta) = H(z, \zeta)(E(z)E(\zeta))^{\frac{1}{2}}$ which is symmetric in the pair of variables $(z, \bar{z})$, $(\zeta, \bar{\zeta})$ ([8], p. 457). Designating by $(E(z))^{\frac{1}{2}}\Psi_m(z)$ the fundamental function for the value λ_m, we have

$$\iint_{\mathfrak{F}^2} E(\zeta)\Psi_k(\zeta)\Psi_p(\zeta)\, d\omega_\zeta = \delta_{kp}, \quad \delta_{kk} = 1, \delta_{kp} = 0, k \neq p. \tag{6.8}$$

Moreover, every function $F(z)$ which can be represented in the form

$$F(z) = \iint_{\mathfrak{F}^2} H(z, \zeta)E(\zeta)T(\zeta)\, d\omega_\zeta \tag{6.9}$$

can also be developed in the uniformly convergent series

$$\sum_{s=1}^{\infty} \frac{2\pi}{\lambda_s} \Psi_s(z) \iint_{\mathfrak{F}^2} E(\zeta)T(\zeta)\Psi_s(\zeta)\, d\omega_\eta, \tag{6.10}$$

([8], pp. 461, 467). Since (6.4) can be written in the form

$$W_n(z) - F_n(z) = \frac{\tau^{(n)}}{2\pi} \iint_{\mathfrak{F}^2} H(z, \zeta)E(\zeta)W_n(\zeta)\, d\omega_\zeta, \tag{6.11}$$

we obtain, therefore, the development

$$W_n(z) - F_n(z) = \sum_{s=1}^{\infty} \frac{2\pi h_s^{(n)}}{\lambda_s} \Psi_s(z), \tag{6.12}$$

$$h_s^{(n)} = \frac{\tau^{(n)}}{2\pi} \iint_{\mathfrak{F}^2} E(\zeta)\Psi_s(\zeta)W_n(\zeta)\, d\omega_\zeta.$$

On substituting (6.12) in (6.4), we obtain

$$h_s^{(n)} = \frac{1}{2\pi} \frac{\lambda_s \tau^{(n)}}{\lambda_s - \tau^{(n)}} \iint_{\mathfrak{F}^2} \Psi_s(\zeta)F_n(\zeta)E(\zeta)\, d\omega_\zeta. \tag{6.13}$$

If $\tau^{(n)}$ converges to λ_m, we obtain, designating $\sum_{s=1}^{m-1} + \sum_{s=m+1}^{\infty}$ by $\sum'$,

$$\left| \sum{}' \frac{h_s^{(n)}}{\lambda_s} \Psi_s(z) \right| = \left| \tau^{(n)} \iint_{\mathfrak{F}^2} F_n(\zeta)(E(\zeta))^{\frac{1}{2}} \sum{}' \frac{\Psi_s(\zeta)\Psi_s(z)(E(\zeta))^{\frac{1}{2}}}{\lambda_s - \tau^{(n)}}\, d\omega_\zeta \right|$$

$$\leqq \tau^{(n)} \left[\iint_{\mathfrak{F}^2} F_n^2(\zeta)E(\zeta)\, d\omega_\zeta \cdot \sum{}' \Psi_s^2(z)(\lambda_s - \tau^{(n)})^{-2} \right]^{\frac{1}{2}} \tag{6.14}$$

$$\leqq C_{20}\tau^{(n)} \left[\iint_{\mathfrak{F}^2} F_n^{2}(\zeta)E(\zeta)\, d\omega_\zeta \sum_{s=1}^{\infty} \frac{\Psi_s^2(z)}{\lambda_s^2} \right]^{\frac{1}{2}},$$

where

$$C_{20} \geqq \max\, [\lambda_s/(\lambda_s - \tau^{(n)})]^2 \qquad (s = 1, 2, \cdots, m-1, m+1, \cdots).$$

The sum

$$\sum_{s=1}^{\infty} \lambda_s^{-2} \Psi_s^2(z)E(z) = \iint_{\mathfrak{F}^2} S_1(z, \zeta)S_1(\zeta, z)\, d\omega_\zeta$$

is finite, since $S_1(z, \zeta)$ is of the same order of infinity as $\log |z - \zeta|$ (see [8], pp. 452, 363). Therefore, by (6.7) and (6.14), we obtain

$$\lim_{n\to\infty} \sum{}' \frac{h_s^{(n)}}{\lambda_s} \Psi_s(z) = 0. \tag{6.15}$$

By (6.15), (6.7), (6.8), and

$$\iint_{\mathfrak{F}^2} \left[\frac{2\pi h_m^{(n)}}{\lambda_m} \Psi_m(z) + \sum{}' \frac{2\pi h_s^{(n)}}{\lambda_s} \Psi_s(z) + F_n(z) \right]^2 E(z)\, d\omega = 1, \tag{6.16}$$

it follows that

$$\lim_{n\to\infty} h_m^{(n)} = \frac{\lambda_m}{2\pi}. \tag{6.17}$$

By (6.15) and (6.17) we get (6.1), (6.12), and (6.7).

7. **Cauchy's problem.** In the case of a hyperbolic equation

$$\mathbf{G}(U) \equiv U_{xy} + A(x, y)U_x + B(x, y)U_y + C(x, y)U = 0 \tag{7.1}$$

we shall determine values of $U(x, y)$ for the domain NOR, limited by the arcs NO, OR, and NR, from the values of U, U_x, U_y along the open curve NR, as given respectively by the functions $\phi(s)$, $\chi_k(s)$ $(k = 1, 2)$. (See Figure 3.) Suppose that NR is a three times differentiable curve which is cut at most once by any line parallel to the axis of x or y. The $\chi_k(s)$ are assumed to possess continuous derivatives of the third order. The parameter s designates the arc-length on NR, measured from R.

We must now determine the $A_k^{(n)}$ in

$$V_n(x, y) = \sum_{\nu=0}^{n} A_\nu^{(n)}\psi_\nu(x, y) \qquad (n = 1, 2, \cdots), \{x, y\} \in NOR \tag{7.2}$$

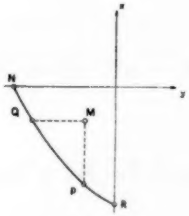

FIG. 3

(where $\psi_\nu(x, y)$ are the functions introduced in (2.13)) in such a way that the V_n converge in NOR to the desired solution. Our method consists in calculating the $A_\nu^{(n)}$ in (7.2) so that

$$\int_{NR}\left[(\phi - V_n)^2 + \left(\chi_1 - \frac{\partial V_n}{\partial x}\right)^2 + \left(\chi_2 - \frac{\partial V_n}{\partial y}\right)^2\right] ds \tag{7.3}$$

shall be a minimum. Designating this minimum by J_n, we want to prove

$$\lim_{n\to\infty} J_n = 0. \tag{7.4}$$

Since A, B, C are analytic functions of x, y, it follows from the hypothesis previously stated that $U(x, y)$ is a twice differentiable function. According to (2.8) we separate U into two parts $U = U_1 + U_2$, where

$$\begin{aligned} U_1(x, y) &= \int_{-1}^{1} \mathrm{E}_1(x, y, t) f\left(x \cdot \frac{1 - t^2}{2}\right)(1 - t^2)^{-\frac{1}{2}}\, dt, \\ U_2(x, y) &= \int_{-1}^{1} \mathrm{E}_2(x, y, t)\, g\left(y \cdot \frac{1 - t^2}{2}\right)(1 - t^2)^{-\frac{1}{2}}\, dt \end{aligned} \tag{7.5}$$

(see (2.8)). It follows from (2.9) that, for $-\frac{1}{2}\alpha < \xi < 0$, $f(\xi)$ and, for $-\frac{1}{2}\beta < \xi < 0$, $g(\xi)$ are functions possessing continuous first derivatives. Therefore,

for the corresponding closed intervals, $f'(\xi) = df(\xi)/d\xi$ and $g'(\xi) = dg(\xi)/d\xi$ are continuous. Here α is the length of the arc NO and β that of the arc RO.

For every ϵ, there exists a polynomial

$$T_n^{(2)}(\xi) = \sum_{k=1}^{n} a_{2k}\xi^k \tag{7.6}$$

so that for $-\frac{1}{2}\alpha \leqq \xi \leqq 0$ we have

$$| f'(\xi) - T_n^{(2)}(\xi) | \leqq \epsilon. \tag{7.7}$$

Let

$$T_n^{(1)}(\xi) = -\int_0^{\xi} T_n^{(2)}(\xi)\, d\xi + f(0). \tag{7.8}$$

Since

$$-\int_0^{\xi} | f'(\xi) - dT_n^{(1)}(\xi)/d\xi | \, d\xi \leqq \epsilon\xi, \tag{7.9}$$

we obtain

$$| f(\xi) - T_n^{(1)}(\xi) | \leqq \tfrac{1}{2}\alpha\epsilon. \tag{7.10}$$

Therefore

$$\begin{aligned} &\left| U_1(x, y) - \sum_{k=0}^{n} a_{2k}\psi_{2k}(x, y) \right| \\ &\leqq \int_{-1}^{1} | \mathrm{E}_1(x, y, t) | \cdot \left| f\left(x \cdot \frac{1 - t^2}{2}\right) - T_n^{(1)}\left(x \cdot \frac{1 - t^2}{2}\right) \right| (1 - t^2)^{-\frac{1}{2}}\, dt \\ &\leqq \int_{-\pi}^{\pi} | \mathrm{E}_1(x, y, \sin\phi) | \cdot | f(\tfrac{1}{2}x\cos^2\phi) - T_n^{(1)}(\tfrac{1}{2}x\cos^2\phi) | \, d\phi \\ &\leqq \epsilon\alpha \, | \mathrm{E}_1 |_{\max}, \end{aligned} \tag{7.11}$$

where $| \mathrm{E}_1 |_{\max} = \max_{\substack{(x,y)\,\epsilon\, NOR \\ -1 \leqq t \leqq 1}} | \mathrm{E}_1(x, y, t) |$, and the $\psi_{2k}(x, y)$ are the functions introduced in (2.13). Furthermore,

$$\begin{aligned} &\left| \frac{\partial U_1(x, y)}{\partial y} - \sum_{k=0}^{n} a_{2k} \frac{\partial \psi_{2k}(x, y)}{\partial y} \right| \\ &\leqq \int_{-1}^{1} \left| \frac{\partial \mathrm{E}_1(x, y, t)}{\partial y} \right| \cdot \left| f\left(x \cdot \frac{1 - t^2}{2}\right) - T_n^{(1)}\left(x \cdot \frac{1 - t^2}{2}\right) \right| (1 - t^2)^{-\frac{1}{2}}\, dt \\ &\leqq \epsilon\alpha \left| \frac{\partial \mathrm{E}_1}{\partial y} \right|_{\max}, \end{aligned} \tag{7.12}$$

and

$$
(7.13)\quad
\begin{aligned}
&\left|\frac{\partial U_1(x, y)}{\partial x} - \sum_{k=0}^{n} a_{2k}\frac{\partial \psi_{2k}(x, y)}{\partial y}\right| \\
&\leqq \int_{-1}^{1}\left|\frac{\partial \mathrm{E}_1(x, y, t)}{\partial x}\right|\cdot\left|f\left(x\cdot\frac{1-t^2}{2}\right) - T_n^{(1)}\left(x\cdot\frac{1-t^2}{2}\right)\right|(1-t^2)^{-\frac{1}{2}}\,dt \\
&\quad + \int_{-1}^{1}|\mathrm{E}_1(x, y, t)|\cdot\left|f'\left(x\cdot\frac{1-t^2}{2}\right) - T_n^{(2)}\left(x\cdot\frac{1-t^2}{2}\right)\right|(1-t^2)^{-\frac{1}{2}}\,dt \\
&\leqq \epsilon\alpha\left|\frac{\partial \mathrm{E}_1}{\partial x}\right|_{\max} + 2\epsilon\,|\mathrm{E}_1|_{\max}.
\end{aligned}
$$

In the same manner, we can approximate U_2, $\partial U_2/\partial x$, $\partial U_2/\partial y$ by an expression $\sum_{k=0}^{n} a_{2k+1}\psi_{2k+1}(x, y)$ and its derivatives.

Therefore, there exists for given functions $\phi(s)$, $\chi_1(s)$, $\chi_2(s)$ possessing continuous derivatives of the third order and for every ϵ an expression $W_n(x, y) = \sum_{k=0}^{n} a_k\psi_k(x, y)$ such that

$$
(7.14)\quad \int_{RN}\left[(\phi - W_n)^2 + \left(\chi_1 - \frac{\partial W_n}{\partial x}\right)^2 + \left(\chi_2 - \frac{\partial W_n}{\partial y}\right)^2\right]ds \leqq C_{21}\epsilon,
$$

where C_{21} is independent of n. Since ϵ can be chosen arbitrarily small, and since $J_n \leqq C_{21}\epsilon$, (7.4) follows.

We shall now prove that in the region NOR we have

$$
(7.15)\quad U(x, y) = \lim_{n\to\infty} V_n(x, y).
$$

Let $H_n(x, y) = U(x, y) - V_n(x, y)$. We have

$$
(7.16)\quad
\begin{aligned}
|H_n(x_P, y_P) - H_n(x_Q, y_Q)| &= \left|\int_Q^P\left[\frac{\partial H_n}{\partial x}dx + \frac{\partial H_n}{\partial y}dy\right]\right| \\
&\leqq \left\{\left[\int_{s_Q}^{s_P}\left(\frac{\partial H_n}{\partial x}\frac{dx}{ds}\right)^2 ds\right]^{\frac{1}{2}} + \left[\int_{s_Q}^{s_P}\left(\frac{\partial H_n}{\partial y}\frac{dy}{ds}\right)^2 ds\right]^{\frac{1}{2}}\right\}\times \mathrm{L}(PQ),
\end{aligned}
$$

where $P(x_P, y_P)$, $Q(x_Q, y_Q)$ are two arbitrary points on the curve NR, and $\mathrm{L}(PQ)$ is the length of the arc PQ. Therefore, $H_n(x, y)$ are functions which, for $(x, y)\,\epsilon\,RN$, are uniformly continuous in n. Hence from $\lim_{n\to\infty}\int_{NR}[U(x, y) - V_n(x, y)]^2\,ds = 0$, we deduce that at every point $(x, y)\,\epsilon\,RN$ we have

$$
(7.17)\quad \lim_{n\to\infty}[U(x, y) - V_n(x, y)] = 0.
$$

It remains to prove that (7.15) holds at every point $M(x_M, y_M)$ of NOR.

Let $H(x, y)$ be an arbitrary solution of (7.1) and $R(x, y; \xi, \eta)$ be the corresponding Riemannian function; i.e., the solution of the adjoint equation

$$R_{\xi\eta} - AR_\xi - BR_\eta + (C - A_\xi - B_\eta)R = 0 \tag{7.18}$$

which for $\xi = x$ and $\eta = y$, respectively, takes on the values

$$R(x, y; x, \eta) = \exp\left[\int_y^\eta A(x, \tau)\, d\tau\right], \quad R(x, y; \xi, y) = \exp\left[\int_x^\xi B(\tau, y)\, d\tau\right]. \tag{7.19}$$

It is well known that ([8], p. 149, (49))

$$\begin{aligned} H(x_M, y_M) &= H(x_Q, y_Q)R(x_M, y_M; x_Q, y_Q) \\ &\quad + \int_Q^P R(x_M, y_M; \xi, \eta)H(\xi, \eta)[B(\xi, \eta)\, d\xi - A(\xi, \eta)\, d\eta] \\ &\quad + \int_Q^P R(x_M, y_M; \xi, \eta)\frac{\partial H(\xi, \eta)}{\partial \xi}\, d\xi + \int_Q^P H(\xi, \eta)\frac{\partial R(x_M, y_M; \xi, \eta)}{\partial \eta}\, d\eta. \end{aligned} \tag{7.20}$$

If we put $H_n(x, y) = U(x, y) - V_n(x, y)$ instead of $H(x, y)$ in (7.20), we obtain

$$\begin{gathered} \lim_{n\to\infty} H_n(x_Q, y_Q) = 0, \qquad \lim_{n\to\infty} \int_{s_Q}^{s_P}\left[\frac{\partial H_n(\xi, \eta)}{\partial \xi}\frac{d\xi}{ds}\right]^2 ds = 0, \\ \lim_{n\to\infty} \int_{s_Q}^{s_P}\left[\frac{\partial H_n(\xi, \eta)}{\partial \eta}\frac{d\eta}{ds}\right]^2 ds = 0. \end{gathered} \tag{7.21}$$

Since $R(x_M, y_M; \xi, \eta)$ is bounded, it follows from (7.20) and (7.21) that

$$\lim_{n\to\infty} H_n(x_M, y_M) = 0; \tag{7.22}$$

i.e., that (7.15) is true in NOR.

8. **Final remarks.** I. Besides the fact that the expressions (4.1), (4.11), and (6.1) converge to the desired solutions, it is possible to give limits, depending upon n, for the minimum of $\max_{z \in \mathfrak{l}^1} |U - W_n|$, etc., by using results concerning the approximation of functions of one variable which have been obtained by S. Bernstein, de la Vallée Poussin, Jackson, Walsh, and others.

II. We notice that it is possible to proceed in an analogous way if mixed data are given instead of boundary values; e.g., if we know the boundary values of the desired function on a part of the boundary and the values of its derivatives on the complementary part. Furthermore, we can apply a similar method to the equation

$$\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2} + a(x, y, z)U = 0.$$

Under suitable hypotheses on $a(x, y, z)$, it can be shown that we have the representation

$$U(x, y, z) = \mathrm{Re} \int_{-1}^{1} \int_{0}^{2\pi} \mathbf{E}(x, y, z, t, \tau) f\left[(x + iy \cos t + iz \sin t) \frac{1 - \tau^2}{2}\right] (1 - \tau^2)^{-\frac{1}{2}} \, dt \, d\tau$$

instead of (2.11), where E is an appropriate function depending only on a. Obviously, the question of the convergence and completeness of the system of particular solutions must be investigated again.

III. The method described in this paper may be used for various problems where numerical results are desired. For such purposes, it is useful to use mechanical or electrical apparatus for the necessary numerical calculations. I shall treat the practical side of this procedure in a later work.

References

1. St. Bergman, *Anwendung eines Koebeschen Satzes auf eine Klasse von Funktionen von zwei komplexen Veränderlichen*, Jahresbericht der deutschen Mathematiker-Vereinigung, vol. 38(1929), pp. 187–190.
2. St. Bergman, *Ueber Kurvenintegrale von Funktionen zweier komplexen Veränderlichen, die die Differentialgleichung $\Delta V + V = 0$ befriedigen*, Mathematische Zeitschrift, vol. 32(1930), pp. 386–406.
3. St. Bergman, *Anwendung des Schwarzschen Lemmas auf eine Klasse von Funktionen von zwei komplexen Veränderlichen*, Jahresbericht der deutschen Mathematiker-Vereinigung, vol. 39(1930), pp. 163–168.
4. St. Bergman, *Ein Näherungsverfahren zur Lösung gewisser partieller linearer Differentialgleichungen*, Zeitschrift für angewandte Mathematik und Mechanik, vol. 11(1931), pp. 323–330.
5. St. Bergman, *Ueber ein Verfahren zur Konstruktion der Näherungslösungen der Gleichung $\Delta u + \tau^2 u = 0$*, Applied Mathematics and Mechanics, vol. 3(1936), pp. 97–107.
6. St. Bergman, *Zur Theorie der Funktionen, die eine lineare partielle Differentialgleichung befriedigen*, Recueil Mathématique, new series, vol. 2(1937), pp. 1169–1198.
7. St. Bergman, *Sur un lien entre la théorie des équations aux derivées partielles elliptiques et celle des fonctions d'une variable complexe*, Comptes Rend. Acad. Sci. Paris, vol. 205(1937), pp. 1198–1200 and pp. 1360–1362.
8. E. Goursat, *Cours d'Analyse Mathématique*, vol. 3, 4th ed., Paris, 1927.
9. R. Hedrick, *Ueber den analytischen Charakter der Lösungen von Differentialgleichungen*, Dissertation, Göttingen, 1901.
10. L. Lichtenstein, *Neuere Entwicklung der Theorie partieller Differentialgleichungen zweiter Ordnung von elliptischem Typus*, Enzyklopädie der mathematischen Wissenschaften, vol. II, C, 12, pp. 1277–1344.
11. A. Sommerfeld, *Randwertaufgaben in der Theorie der partiellen Differentialgleichungen*, Enzyklopädie der mathematischen Wissenschaften, vol. II, A, 7c, pp. 504–570.
12. S. Warschawski, *Ueber einen Satz von O. D. Kellogg*, Nachrichten der Gesellschaft der Wissenschaften zu Göttingen, 1932, pp. 73–86.

Massachusetts Institute of Technology.

A COMPATIBLE INTEGRO-DIFFERENTIAL SYSTEM

By Lowell S. Winton

1. **Introduction.** This paper is concerned with the integro-differential system

$$(1.1)\qquad L^i[u(x:s)] = \frac{\partial u^i(x:s)}{\partial x} + \Phi^i_j(x:s)u^j(x:s) + \int_\alpha^\beta K^i_j(x:s,t)u^j(x:t)\,dt,$$

$$(1.2)\qquad \begin{aligned} U^i[u(x:s)] &= \alpha^i_j(s)u^j(a:s) + \beta^i_j(s)u^j(b:s) \\ &\quad + \int_\alpha^\beta [A^i_j(s,r)u^j(a:r) + B^i_j(s,r)u^j(b:r)]\,dr. \end{aligned}$$

$\Phi^i_j(x:s)$, $K^i_j(x:s,t)$, $\alpha^i_j(s)$, $\beta^i_j(s)$, $A^i_j(s,r)$, and $B^i_j(s,r)$ are to be known and continuous in the regions I_{xs} $[a \leqq x \leqq b,\ \alpha \leqq s \leqq \beta]$, I_{xst} $[a \leqq x \leqq b,\ \alpha \leqq s,\ t \leqq \beta]$, J_s $[\alpha \leqq s \leqq \beta]$, and J_{sr} $[\alpha \leqq s,\ r \leqq \beta]$. We seek a set $u^i(x:s)$, continuous and with continuous first partial x derivatives in I_{xs} satisfying the non-homogeneous system

$$(1.3)\qquad L^i[u(x:s)] = f^i(x:s), \qquad U^i[u(x:s)] = X^i(s),$$

under the hypothesis that the homogeneous system

$$(1.4)\qquad L^i[u(x:s)] = 0, \qquad U^i[u(x:s)] = 0$$

is compatible of order k.

M. T. Hu[1] studied the incompatible case for a single equation and one boundary condition. §§2, 3, 4, and 5 consist largely of an extension to systems of equations of the results of his paper that are essential here. Since for the most part such generalizations are simple, we shall simply state results, elaborating only where the work is new or different.

Our main concern will be to devise generalized Green's functions $H^i_j(x, y:s)$ and $G^i_j(x, y:s, t)$ analogous to those for compatible differential systems developed by W. W. Elliott.[2] In addition we shall devise auxiliary functions $P^i_j(x:s)$ and $Q^i_j(x:s, t)$ such that the solution of (1.3), when it exists, may be written in the form

$$(1.5)\qquad \begin{aligned} u^i(x:s) &= \int_a^b H^i_j(x,y:s)f^j(y:s)\,dy + \int_a^b\int_\alpha^\beta G^i_j(x,y:s,t)f^j(y:t)\,dy\,dt \\ &\quad + P^i_j(x:s)X^j(s) + \int_\alpha^\beta Q^i_j(x:s,t)X^j(t)\,dt + A^s u^i_s(x:s). \end{aligned}$$

Received April 18, 1939.

[1] M. T. Hu, *Linear integro-differential equations with a boundary condition*, Transactions of the American Mathematical Society, vol. 19(1918), pp. 363–407.

[2] W. W. Elliott, *Generalized Green's functions for compatible differential systems*, American Journal of Mathematics, vol. 50(1928), pp. 243–258.

In this the A^δ are arbitrary constants and the $u_\delta^i(x:s)$ are the k linearly independent solutions of (1.4). These results are given in §§6 and 7.

Throughout, the indices will have the following ranges: $c, d, i, j = 1, \cdots, n$; $I, J = 1, \cdots, 2n$; $\gamma, \delta = 1, \cdots, k$.

2. The integro-differential equations.

THEOREM 2.1. *The set of integro-differential equations*

$$L^i[u(x:s)] = f^i(x:s)$$

has a unique set of solutions reducing to a set of assigned initial functions $u^i(y:s)$ at some point y in the interval (a, b). This solution is given by

$$u^i(x:s) = R_j^i(x, y:s)u^j(y:s) + \int_\alpha^\beta S_j^i(x, y:s, t)u^j(y:t)\,dt + \omega^i(x, y:s). \tag{2.1}$$

In (2.1)

$$\omega^i(x, y:s) = \int_y^x \left[R_j^i(x, \xi:s)f^j(\xi:s) + \int_\alpha^\beta S_j^i(x, \xi:s, t)f^j(\xi:t)\,dt\right]d\xi,$$

and the functions $R_j^i(x, y: s)$ and $S_j^i(x, y: s, t)$ are defined as follows. Let $r_j^i(x:s)$ be, for each i, a set of functions, continuous and with continuous x derivatives in I_{xs}, such that

$$\frac{\partial r_j^i(x:s)}{\partial x} - \Phi_j^c(x:s)r_c^i(x:s) = 0, \tag{2.2}$$

and let them be so chosen that

$$W(x:s) = | r_j^i(x:s) | \neq 0 \qquad \text{in } I_{xs}. \tag{2.3}$$

Let

$$\rho_i^j(x:s) = \frac{\text{co-factor of } r_j^i(x:s) \text{ in } W(x:s)}{W(x:s)}, \tag{2.4}$$

then

$$R_j^i(x, y:s) = \rho_c^i(x:s)r_j^c(y:s). \tag{2.5}$$

Also, clearly

$$R_j^i(y, y:s) = \delta_j^i. \tag{2.6}$$

By differentiating the identity

$$r_j^c(x:s)\rho_c^i(x:s) = \delta_j^i$$

with respect to x, it may be shown that the $\rho_j^i(x:s)$ and thus, from (2.5), that the $R_j^i(x, y:s)$, for each j, satisfy

$$\frac{\partial R_j^i(x, y:s)}{\partial x} + \Phi_c^i(x:s)R_j^c(x, y:s) = 0. \tag{2.7}$$

Again from (2.5)

$$\frac{\partial R^i_j(x, y:s)}{\partial y} - \Phi^c_j(y:s)R^i_c(x, y:s) = 0. \tag{2.8}$$

In the system of Volterra integral equations

$$u^i(x:s) = F^i(x:s) + \int_y^x \int_\alpha^\beta [-R^i_c(x, \xi:s)K^c_j(\xi:s, t)]u^j(\xi:t)\, d\xi\, dt \tag{2.9}$$

let $\Theta^i_j(x, \xi:s, t)$ be the resolvent system to the kernel system $[-R^i_c(x, \xi:s)K^c_j(\xi, s:t)]$; then

$$S^i_j(x, y:s, t) = \int_y^x \Theta^i_c(x, \xi:s, t)R^c_j(\xi, y:t)\, d\xi. \tag{2.10}$$

Also we note that

$$S^i_j(y, y:s, t) = 0. \tag{2.11}$$

3. **The boundary problem.** As in Hu's paper the determination of the $u^i(y:s)$ in (2.1) so that the $u^i(x:s)$ will satisfy

$$U^i[u(x:s)] = X^i(s)$$

leads to the system of Fredholm integral equations

$$u^i(y:s) = \Gamma^i_j(y:s)F^j(y:s) + \int_\alpha^\beta P^i_j(y:s, t)u^j(y:t)\, dt, \tag{3.1}$$

in which the $\Gamma^i_j(y:s)$, $F^i(y:s)$, and $P^i_j(y:s, t)$ are defined as follows. If

$$\begin{aligned} g^i_j(y:s) &= \alpha^i_c(s)R^c_j(a, y:s) + \beta^i_c(s)R^c_j(b, y:s),\\ h^i_j(y:s, t) &= -\{A^i_c(s, t)R^c_j(a, y:t) + B^i_c(s, t)R^c_j(b, y:t) + U^i_{xs}[S_j(x, y:s, t)]\},\\ F^i(y:s) &= X^i(s) - U^i[\omega(x, y:s)], \end{aligned} \tag{3.2}$$

and if

$$g(y:s) = |\, g^i_j(y:s)\,| \neq 0 \qquad \text{in } I_{ys}, \tag{3.3}$$

then

$$\Gamma^j_i(y:s) = \frac{\text{co-factor of } g^i_j(y:s) \text{ in } g(y:s)}{g(y:s)}$$

and

$$P^i_j(y:s, t) = \Gamma^i_c(y:s)h^c_j(y:s, t).$$

The condition (3.3) is the same as assuming that the system

$$\begin{aligned} l^i[u(x:s)] &= \frac{\partial u^i(x:s)}{\partial x} + \Phi^i_j(x:s)u^j(x:s) = 0,\\ W^i[u(x:s)] &= \alpha^i_j(s)u^j(a:s) + \beta^i_j(s)u^j(b:s) = 0 \end{aligned} \tag{3.4}$$

is incompatible for all s in J_s. It is likewise independent of the initial point y, as it can be shown that

$$g(y:s) = M(s)W(y:s),$$

and we assume throughout that $W(y:s) \neq 0$ in I_{ys}.

The order of compatibility of (1.4) is thus that of the homogeneous set of Fredholm integral equations

$$u^i(y:s) = \int_\alpha^\beta P_j^i(y:s,\, t)u^j(y:t)\,dt, \tag{3.5}$$

and the application of (2.1) will show that the Fredholm determinant[3] $D(y)$ of the kernel system $P_j^i(y:s,\, t)$ vanishes everywhere or nowhere in I_y.

If (3.5) is compatible of order k and has the linearly independent solutions $u_k^i(y:s)$, and if $\varphi_i^k(y:s)$ denote the k linearly independent solutions of

$$\varphi_i(y:s) = \int_\alpha^\beta P_i^j(y:t,\, s)\varphi_j(y:t)\,dt, \tag{3.6}$$

the adjoint set of equations to (3.5), we may state the following theorem.

THEOREM 3.1. *A necessary and sufficient condition that* (3.1) *have a solution when* (3.5) *is compatible of order k is that the $F^i(y:s)$ satisfy the conditions*

$$\int_\alpha^\beta \varphi_i^k(y:s)\Gamma_c^i(y:s)F^c(y:s)\,ds = 0; \tag{3.7}$$

and, if these conditions are satisfied, the solution is given by

$$u^i(y:s) = \Gamma_c^i(y:s)F^c(y:s) + \int_\alpha^\beta q_c^i(y:s,\, t)\Gamma_j^c(y:t)F^j(y:t)\,dt + A^k u_k^i(y:s), \tag{3.8}$$

where the $q_c^i(y:s,\, t)$ form a pseudo-resolvent system to the kernel system $P_c^i(y:s,\, t)$.[4]

The condition (3.7) also serves as a necessary and sufficient condition that (1.3) have a solution when (1.4) is compatible of order k, and another more useful form of this condition will be given later.

4. **Integro-linear dependence.** With Hu we say that the n $U^i[u]$ given by (1.2), involving the $2n$ arbitrary continuous functions $u^j(a:s)$ and $u^j(b:s)$, will be said to be self integro-linearly dependent if there exists a set of functions $C_i(s)$, continuous in J_s, and not all zero, such that

$$\int_\alpha^\beta C_i(s)U^i[u(x:s)]\,ds = 0 \tag{4.1}$$

[3] Charles Platrier, *Sur les mineurs de la fonction déterminante de Fredholm et sur les systèmes d'équations intégrales linéaires*, Journal de Mathématiques pures et appliquées, (6), vol. 9(1913), pp. 233–304.

[4] This follows from an extension to systems of Fredholm equations of a paper by W. A. Hurwitz, *On the pseudo-resolvent to the kernel of an integral equation*, Transactions of the American Mathematical Society, vol. 13(1912), pp. 405–418.

for all $u^j(a:s)$ and $u^j(b:s)$. If no such functions exist, the $U^i[u]$ will be said to be self integro-linearly independent. If in (4.1) the coefficients of the $2n$ arbitrary functions $u^j(a:s)$ and $u^j(b:s)$ are set equal to zero, we find a necessary and sufficient condition for the self integro-linear independence of the $U^i[u]$; namely, that the set of $2n$ equations

$$\begin{aligned} C_j(s)\alpha_i^j(s) + \int_\alpha^\beta C_j(r)A_i^j(r, s)\, dr &= 0, \\ C_j(s)\beta_i^j(s) + \int_\alpha^\beta C_j(r)B_i^j(r, s)\, dr &= 0 \end{aligned} \tag{4.2}$$

in the $C_i(s)$ have no set of non-trivial solutions.

THEOREM 4.1. *If the $U^i[u]$ are self integro-linearly dependent, the system (1.4) is compatible.*

For if such a set $C_i(s)$ exists, the set of functions

$$\varphi_i(y:s) = C_j(s)g_i^j(y:s)$$

satisfies (3.6), as may be verified by direct substitution.

When there are $2n$ expressions of the form of (1.2), they will be integro-linearly independent if there exist no non-zero $C_I(s)$ such that

$$\int_\alpha^\beta C_I(s)U^I[u(x:s)]\, ds = 0; \tag{4.3}$$

that is, if the $2n$ equations

$$\begin{aligned} C_J(s)\alpha_i^J(s) + \int_\alpha^\beta C_J(r)A_i^J(r, s)\, dr &= 0, \\ C_J(s)\beta_i^J(s) + \int_\alpha^\beta C_J(r)B_i^J(r, s)\, dr &= 0 \end{aligned} \tag{4.4}$$

have no non-trivial solution. If

$$a(s) = \begin{vmatrix} \alpha_1^1(s) & \cdots & \alpha_1^{2n}(s) \\ \cdots & \cdots & \cdots \\ \beta_n^1(s) & \cdots & \beta_n^{2n}(s) \end{vmatrix} \neq 0 \quad \text{in } J_s, \tag{4.5}$$

the functions

$$\begin{aligned} a_J^i(s) &= \frac{\text{co-factor of } \alpha_i^J(s) \text{ in } a(s)}{a(s)}, \\ a_J^{n+i}(s) &= \frac{\text{co-factor of } \beta_i^J(s) \text{ in } a(s)}{a(s)} \end{aligned} \tag{4.6}$$

will exist. (4.4) may then be written in the form

$$C_I(s) + \int_\alpha^\beta C_J(r)[A_j^J(r, s)a_I^j(s) + B_j^J(r, s)a_I^{n+j}(s)]\, dr = 0. \tag{4.7}$$

Hence we have the following theorem.

THEOREM 4.2. *A necessary and sufficient condition that the $U^I[u]$ be integro-linearly independent, subject to* (4.5), *is that the Fredholm determinant*[5] *to the kernel system of the equations* (4.7) *does not vanish.*

Also it is clear from the definition of integro-linear dependence that a necessary condition for the integro-linear independence of the $U^i[u]$ and the $U^{i+n}[u]$ is that both sets be self integro-linearly independent.

It is important to show that, given the self integro-linearly independent set $U^i[u]$, there can be found a self integro-linearly independent set $U^{i+n}[u]$ such that the combined set $U^I[u]$ will be integro-linearly independent. As Hu was unable to give a general proof for this fact, we give the work in a somewhat detailed form. To show this fact we shall have need of

THEOREM 4.3. *If the $U^i[u]$ are self integro-linearly independent, and if not all the n-rowed determinants in*

$$
(4.8) \qquad \begin{vmatrix} \alpha_1^1(s) & \cdots & \alpha_n^1(s) & \beta_1^1(s) & \cdots & \beta_n^1(s) \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \alpha_1^n(s) & \cdots & \alpha_n^n(s) & \beta_1^n(s) & \cdots & \beta_n^n(s) \end{vmatrix}
$$

vanish together in J_s, there exist no non-zero $C_i(s)$ such that

$$
(4.9) \qquad C_i(s) + \int_\alpha^\beta C_e(r)[A_j^e(r, s)a_i^j(s) + B_j^e(r, s)a_i^{n+j}(s)]\, dr = 0,
$$

$$
(4.10) \qquad \int_\alpha^\beta C_e(r)[A_j^e(r, s)a_{n+i}^j(s) + B_j^e(r, s)a_{n+i}^{n+j}(s)]\, dr = 0.
$$

The non-vanishing of the n-rowed determinants in (4.8) is not a new restriction as it is a necessary condition for (3.3) to hold. It enables[6] us to choose the $\alpha_j^{n+i}(s)$ and $\beta_j^{n+i}(s)$ so that (4.5) is satisfied and thus so that the functions (4.6) exist. If (4.9) is multiplied by $\alpha_j^i(s)$, (4.10) by $\alpha_j^{i+n}(s)$, and the results are added, we have the first of equations (4.2). Similar use of $\beta_j^i(s)$ and $\beta_j^{i+n}(s)$ yields the second of equations (4.2). Under our hypothesis (4.2) has no non-trivial solution.

THEOREM 4.4. *If the $U^i[u]$ satisfy the conditions of Theorem* 4.3, *there can be found self integro-linearly independent $U^{i+n}[u]$ such that the combined set $U^I[u]$ will be integro-linearly independent.*

To show this we first write (4.7) in the form

$$
(4.11) \qquad \begin{aligned} C_i(s) &= \int_\alpha^\beta [C_j(r)K_i^j(r, s) + C_{n+j}(r)K_i^{n+j}(r, s)]\, dr, \\ C_{n+i}(s) &= \int_\alpha^\beta [C_j(r)K_{n+i}^j(r, s) + C_{n+j}(r)K_{n+i}^{n+j}(r, s)]\, dr, \end{aligned}
$$

[5] Platrier, loc. cit.

[6] G. A. Bliss, *The problem of Mayer with variable end points*, Transactions of the American Mathematical Society, vol. 19(1918), p. 312, auxiliary theorem.

in which

$$
(4.12)\qquad \begin{aligned} K_i^j(r, s) &= -A_c^j(r, s)a_i^c(s) - B_c^j(r, s)a_i^{n+c}(s), \\ K_{n+i}^j(r, s) &= -A_c^j(r, s)a_{n+i}^c(s) - B_c^j(r, s)a_{n+i}^{n+c}(s), \end{aligned}
$$

and

$$
(4.13)\qquad \begin{aligned} K_i^{n+j}(r, s) &= -A_c^{n+j}(r, s)a_i^c(s) - B_c^{n+j}(r, s)a_i^{n+c}(s), \\ K_{n+i}^{n+j}(r, s) &= -A_c^{n+j}(r, s)a_{n+i}^c(s) - B_c^{n+j}(r, s)a_{n+i}^{n+c}(s). \end{aligned}
$$

Under the stated conditions we can choose the $\alpha_j^{n+i}(s)$ and $\beta_j^{n+i}(s)$ so that (4.5) is satisfied and thus so that the $A_j^{n+i}(r, s)$ and $B_j^{n+i}(r, s)$ are uniquely determinable in terms of the $K_j^{n+i}(r, s)$ and $K_{n+j}^{n+i}(r, s)$. Choose

$$
(4.14)\qquad K_{n+j}^{n+i}(r, s) = 0, \qquad K_j^{n+i}(r, s) = \psi_\gamma^j(s)\sigma_\gamma^i(r).
$$

In this the $\psi_\gamma^i(s)$ are the k linearly independent solutions of

$$
(4.15)\qquad \psi^i(s) = \int_\alpha^\beta K_c^i(s, r)\psi^c(r)\,dr,
$$

and are to form a normalized orthogonal set. The $\sigma_\gamma^j(s)$ are determined as follows. Let the k linearly independent solutions of

$$
(4.16)\qquad \varphi_i(s) = \int_\alpha^\beta K_i^c(r, s)\varphi_c(r)\,dr
$$

be $\varphi_i^\gamma(s)$. Then by Theorem 4.3 the sets

$$
(4.17)\qquad \xi_i^\gamma(s) = \int_\alpha^\beta \varphi_j^\gamma(r)K_{n+i}^j(r, s)\,dr
$$

are linearly independent. There thus exists at least one set $\sigma_\gamma^i(s)$ that is biorthogonal to the $\xi_i^\gamma(s)$; that is, such that

$$
(4.18)\qquad \int_\alpha^\beta \sigma_\gamma^i(s)\xi_i^\delta(s)\,ds = \delta_\gamma^\delta.
$$

The choice (4.14) reduces (4.11) to the form

$$
(4.19)\qquad \begin{aligned} C_i(s) &= \int_\alpha^\beta [C_j(r)K_i^j(r, s) + C_{n+j}(r)\sigma_\gamma^j(r)\psi_\gamma^i(s)]\,dr, \\ C_{n+i}(s) &= \int_\alpha^\beta C_j(r)K_{n+i}^j(r, s)\,dr. \end{aligned}
$$

Thus

$$
(4.20)\qquad C_i(s) = \int_\alpha^\beta C_j(r)K_i^j(r, s)\,dr + \int_\alpha^\beta \int_\alpha^\beta C_j(r)K_{n+c}^j(r, t)\sigma_\gamma^c(t)\psi_\gamma^i(s)\,dr\,dt.
$$

Multiply (4.20) by $\psi_s^i(s)$ and integrate. Since the $\psi_s^i(s)$ are normal orthogonal, we have

$$\int_\alpha^\beta \int_\alpha^\beta C_\alpha(r) K_{n+j}^\alpha(r, t)\sigma_\gamma^j(t)\, dr\, dt = 0. \tag{4.21}$$

Hence

$$C_i(s) = K_\gamma \varphi_i^\gamma(s).$$

The K_γ may then be evaluated by means of (4.21) and with the aid of (4.18) they are found to be zero. Thus by (4.19) the $C_l(s)$ are all zero for this choice of $K_i^{n+j}(r, s)$ and $K_{n+i}^{n+j}(r, s)$, and thus the $A_i^{n+j}(r, s)$ and $B_i^{n+j}(r, s)$ determined by (4.13) will give us the desired result.

5. **The adjoint system.** The set of integro-differential expressions

$$M_i[v(x:s)] = -\frac{\partial v_i(x:s)}{\partial x} + \Phi_i^j(x:s)v_j(x:s) + \int_\alpha^\beta K_i^j(x:t, s)v_j(x:t)\, dt \tag{5.1}$$

will be called the set adjoint to (1.1). For any two points x_1 and x_2 and for any two sets of functions $u^i(x:s)$ and $v_i(x:s)$ the following identity then holds.

$$\begin{aligned}\int_{x_1}^{x_2}\int_\alpha^\beta \{v_i(x:s)L^i[u(x:s)] - u^i(x:s)M_i[v(x:s)]\}\, ds \\ = \int_\alpha^\beta [u^i(x_2:s)v_i(x_2:s) - u^i(x_1:s)v_i(x_1:s)]\, ds.\end{aligned} \tag{5.2}$$

We have also the symmetrical relationships

$$R_j^i(x, y:s) = \bar{R}_j^i(y, x:s), \qquad S_j^i(x, y:s, t) = \bar{S}_j^i(y, x:t, s),$$

where the bar indicates the corresponding function for the adjoint.

The boundary conditions that are to be joined to (5.1) are of the form

$$\begin{aligned}V_i[v] = \gamma_i^j(s)v_j(a:s) + \delta_i^j(s)v_j(b:s) \\ + \int_\alpha^\beta [C_i^j(s, r)v_j(a:r) + D_i^j(s, r)v_j(b:r)]\, dr.\end{aligned} \tag{5.3}$$

They will be so determined that for all $u^j(a:s)$, $u^j(b:s)$, $v_j(a:s)$, and $v_j(b:s)$,

$$\begin{aligned}\int_\alpha^\beta [u^i(b:s)v_i(b:s) - u^i(a:s)v_i(b:s)]\, ds \\ = \int_\alpha^\beta [U^i[u]V_{2n-i+1}[v] + U^{2n-i+1}[u]V_i[v]]\, ds\end{aligned} \tag{5.4}$$

in which the $V_{2n-i+1}[v]$ are n other expressions of the form (5.3). If the methods of Hu's paper are then followed, the following theorems can be proved.

THEOREM 5.1. *For every two self integro-linearly independent sets of expressions $U^i[u]$ and $U^{2n-i+1}[u]$ such that the combined set is integro-linearly independent and satisfies* (4.5) *there exist two unique sets of expressions $V_i[v]$ and $V_{2n-i+1}[v]$ which are integro-linearly independent and satisfy* (4.5), *and which with the $U^I[u]$ satisfy* (5.4) *for all $u^i(a:s)$, $u^i(b:s)$, $v_i(a:s)$, and $v_i(b:s)$.*

From (5.2) and (5.4) we can then write Green's formula

$$\int_a^b \int_\alpha^\beta \{v_i(x:s)L^i[u(x:s)] - u^i(x:s)M_i[v(x:s)]\}\, ds = \int_\alpha^\beta \{U^i[u]V_{2n-i+1}[v] + U^{2n-i+1}[u]V_i[v]\}\, ds. \tag{5.5}$$

THEOREM 5.2. *The $V_i[v]$ are determined except for integro-linear combinations when the $U^i[u]$ are given.*

That is, for any two choices $U_1^{2n-i+1}[u]$ and $U_2^{2n-i+1}[u]$ the following relations hold.

$$\begin{aligned} V_I^1[v] &= V_I^2[v] + M_I^{1j}(s)V_j^2[v] + \int_\alpha^\beta N_I^{1j}(s, r)V_j^2[v]\, dr, \\ V_I^2[v] &= V_I^1[v] + M_I^{2j}(s)V_j^1[v] + \int_\alpha^\beta N_I^{2j}(s, r)V_j^1[v]\, dr. \end{aligned} \tag{5.6}$$

From these equations it is clear that the following theorem is true.

THEOREM 5.3. *If $v_i(x:s)$ is any set of functions satisfying $V_i[v] = 0$, then for any two choices of $U^{2n-i+1}[u]$ we have*

$$V_{2n-i+1}^1[v] = V_{2n-i+1}^2[v].$$

Again by Hu's methods we can prove

THEOREM 5.4. *If there exist k linearly independent sets of functions $v_i^\gamma(x:s)$ such that $V_i[v^\gamma] = 0$, then the $V_{2n-i+1}[v^\gamma]$ form linearly independent sets.*

For purposes of notation we write $V_{2n-i+1}[v^\gamma] = V_i^\gamma(s)$.

The system

$$M_i[v(x:s)] = 0, \qquad V_i[v(x:s)] = 0 \tag{5.7}$$

will be called the system adjoint to (1.4). Its k linearly independent solutions will be written $v_i^\gamma(x:s)$ when (1.4) is compatible of order k, this being consistent with the following theorem.

THEOREM 5.5. *The index of compatibility of* (5.7) *is the same as that of* (1.4).

It can be verified that k linearly independent solutions of (3.6) are given by

$$\varphi_i^\gamma(y:s) = V_j^\gamma(s)g_i^j(y:s).$$

If these values are substituted in (3.7) and the result simplified by (5.5), we find a more useful form of the condition that (1.3) have a solution when (1.4) is compatible of order k.

THEOREM 5.6. *A necessary and sufficient condition that* (1.3) *have a solution when* (1.4) *is compatible of order k is that the $f^i(x:s)$ and $X^i(s)$ be such that*

$$\int_a^b \int_\alpha^\beta f^i(x:s)v_i^\gamma(x:s)\,dx\,ds = \int_\alpha^\beta X^i(s)V_i^\gamma(s)\,ds. \tag{5.8}$$

COROLLARY 5.1. *If* (1.4) *is compatible of order k, the semi-homogeneous systems*

$$L^i[u(x:s)] = f^i(x:s), \qquad U^i[u(x:s)] = 0 \tag{5.9}$$

and

$$L^i[u(x:s)] = 0, \qquad U^i[u(x:s)] = X^i(s) \tag{5.10}$$

will have solutions if and only if the $f^i(x:s)$ and $X^i(s)$ satisfy the respective conditions

$$\int_a^b \int_\alpha^\beta f^i(x:s)v_i^\gamma(x:s)\,dx\,ds = 0, \tag{5.11}$$

and

$$\int_\alpha^\beta X^i(s)V_i^\gamma(s)\,ds = 0. \tag{5.12}$$

We now state as a corollary to the corresponding lemma in Hu's paper the following lemma which is used throughout this paper.

LEMMA I. *If $d_j^i(s)$ and $D_j^i(s, t)$ are continuous functions such that*

$$d_c^i(s)\varphi^c(s) + \int_\alpha^\beta D_c^i(s, t)\varphi^c(t)\,dt = 0$$

for every arbitrary, continuous set of functions $\varphi^i(s)$, then

$$d_c^i(s) = 0; \qquad D_c^i(s, t) = 0.$$

6. **The generalized Green's function.** We seek two sets of functions $H_j^i(x, y:s)$ and $G_j^i(x, y:s, t)$, independent of the $f^i(x:s)$ in the semi-homogeneous system (5.9), such that the solution of that system may be written in the form

$$\begin{aligned} u^i(x:s) = {} & \int_a^b H_j^i(x, y:s)f^j(y:s)\,dy \\ & + \int_a^b \int_\alpha^\beta G_j^i(x, y:s, t)f^j(y:t)\,dy\,dt + A^\gamma u_\gamma^i(x:s), \end{aligned} \tag{6.1}$$

whenever (5.11) is satisfied. We complete our definition of the generalized Green's functions with the following requirements. $H_j^i(x, y:s)$ is to be continuous throughout the regions

$$I_1\,[a \leqq x < y;\ \alpha \leqq s \leqq \beta] \quad \text{and} \quad I_2\,[y < x \leqq b;\ \alpha \leqq s \leqq \beta] \tag{6.2}$$

and possess continuous first partial x derivatives in those regions. At $x = y$ they are to have a discontinuity of the type

$$H_j^i(y^+, y:s) - H_j^i(y^-, y:s) = \delta_j^i\,. \tag{6.3}$$

The $G^i_j(x, y:s, t)$ are to be continuous throughout I_{xyst} and are to have continuous first partial x derivatives throughout I_1 and I_2.

THEOREM 6.1. *Generalized Green's functions for the compatible system* (1.4) *exist and for each j they satisfy*

$$(6.4)\qquad \begin{aligned} &\frac{\partial}{\partial x} H^i_j(x, y:s) + \Phi^i_c(x:s)H^c_j(x, y:s) = 0,\\ &\alpha^i_c(s)H^c_j(a, y:s) + \beta^i_c(s)H^c_j(b, y:s) = 0\end{aligned}$$

and

$$(6.5)\qquad \begin{aligned} L^i_{xs}[G_j(x, y:s, t)] &= -K^i_c(x:s, t)H^c_j(x, y:t) + \psi^i_\gamma(x:s)v^\gamma_j(y:t),\\ U^i_{xs}[G_j(x, y:s, t)] &= -A^i_c(s, t)H^c_j(a, y:t) - B^i_c(s, t)H^c_j(b, y:t) + \rho^i_\gamma(s)v^\gamma_j(y:t),\end{aligned}$$

the $\psi^i_\gamma(x:s)$ and $\rho^i_\gamma(s)$ being subject to the conditions

$$(6.6)\qquad \int_a^b\int_\alpha^\beta v^\delta_i(x:s)\psi^i_\gamma(x:s)\,dx\,ds - \int_\alpha^\beta V^\delta_i(s)\rho^i_\gamma(s)\,ds + \delta^\delta_\gamma = 0.$$

A set of functions satisfying (5.11) is given by

$$(6.7)\qquad f^i(x:s) = \begin{vmatrix} g^i(x:s) & v^1_i(x:s) & \cdots & v^k_i(x:s)\\ g^c v^1_c & v^1_c v^1_c & \cdots & v^1_c v^k_c\\ \cdots & \cdots & \cdots & \cdots\\ g^c v^k_c & v^k_c v^1_c & \cdots & v^k_c v^k_c\end{vmatrix}$$

for arbitrary continuous $g^i(x:s)$ and in which

$$(6.8)\qquad \begin{aligned} g^c v^\gamma_c &= \int_a^b\int_\alpha^\beta g^c(x:s)v^\gamma_c(x:s)\,dx\,ds,\\ v^\delta_c v^\gamma_c &= \int_a^b\int_\alpha^\beta v^\delta_c(x:s)v^\gamma_c(x:s)\,dx\,ds.\end{aligned}$$

In (6.1) let the $f^i(x:s)$ be given by (6.7) and apply the L^i. It is found that if (6.1) is to be a solution of (5.9),

$$(6.9)\qquad \int_a^b A^i_j(x, y:s)f^j(y:s)\,dy + \int_a^b\int_\alpha^\beta B^i_j(x, y:s, t)f^j(y:t)\,dy\,dt = 0,$$

in which

$$(6.10)\qquad \begin{aligned} A^i_j(x, y:s) &= \frac{\partial}{\partial x} H^i_j(x, y:s) + \Phi^i_c(x:s)H^c_j(x, y:s) = 0,\\ B^i_j(x, y:s, t) &= L^i_{xs}[G_j(x, y:s, t)] + K^i_c(x:s, t)H^c_j(x, y:t).\end{aligned}$$

(6.9) may be written in the form

$$(6.11)\qquad |v^\gamma_c v^\delta_c| \int_\alpha^\beta A^i_c(x, y:s)g^c(y:s)\,dy + \int_a^b\int_\alpha^\beta D^i_c(x, y:s, t)g^c(y:t)\,dy\,dt = 0,$$

in which

$$
(6.12) \qquad D^i_j(x, y:s, t) = \begin{vmatrix} B^i_j(x, y:s, t) & P^i_1(x:s) & \cdots & P^i_k(x:s) \\ v^1_j(y:t) & v^1_c v^1_c & \cdots & v^1_c v^k_c \\ \cdots & \cdots & \cdots & \cdots \\ v^k_j(y:t) & v^k_c v^1_c & \cdots & v^k_c v^k_c \end{vmatrix} = 0,
$$

if Lemma I is applied, where

$$
P^i_\gamma(x:s) = \int_a^b A^i_c(x, y:s) v^\gamma_c(y:s)\, dy + \int_a^b \int_\alpha^\beta B^i_c(x, y:s, t) v^\gamma_c(y:t)\, dy\, dt.
$$

Since the $g^i(x:s)$ are arbitrary and the $| v^\gamma_c v^\delta_c |$, the Grammian of the sets of functions $v^\alpha_i(x:s)$, does not vanish, it is necessary that[7]

$$
B^i_j(x, y:s, t) = \psi^i_\gamma(x:s) v^\gamma_j(y:t),
$$

the $\psi^i_\gamma(x:s)$ being k sets of arbitrary continuous functions. Similarly, if the U^i are applied, the second conditions in (6.4) and (6.5) will be found necessary, the $\rho^i_\gamma(s)$ being arbitrary.

It can readily be verified that a set of functions satisfying (6.3) and (6.4) are given by

$$
(6.13) \qquad \begin{aligned} H^i_j(x, y:s) &= -R^i_c(x, y:s)\Gamma^c_d(y:s)\beta^d_e(s)R^e_j(b, y:s), && x < y, \\ H^i_j(x, y:s) &= R^i_c(x, y:s)\Gamma^c_d(y:s)\alpha^d_e(s)R^e_j(a, y:s), && x > y, \end{aligned}
$$

whenever the condition (3.3) is satisfied. As these are the ordinary Green's functions[8] for the incompatible system (3.4), these functions are unique. If (3.3) is not satisfied, these functions will not exist and thus (3.3) is a necessary condition for the solution of (5.9) in the form (6.1).

By (5.8), (6.5) will have a solution if and only if

$$
(6.14) \qquad \begin{aligned} &\int_a^b \int_\alpha^\beta v^\gamma_i(x:s)[-K^i_c(x:s, t)H^c_j(x, y:t) + \psi^i_\delta(x:s)v^\delta_j(y:t)]\, dx\, ds \\ &= \int_\alpha^\beta V^\gamma_i(s)[-A^i_c(s, t)H^c_j(a, y:t) - B^i_c(s, t)H^c_j(b, y:t) + \rho^i_\delta(s)v^\delta_j(y:t)]\, ds. \end{aligned}
$$

As an aid in simplifying this last we note that for arbitrary continuous $f^j(t)$

$$
\begin{aligned} &\int_\alpha^\beta \int_\alpha^\beta V_{2n-i+1}[v^\gamma(x:s)][-A^i_c(s, t)H^c_j(a, y:t) - B^i_c(s, t)H^c_j(b, y:t)]f^j(t)\, dt\, ds \\ &\qquad = -\int_\alpha^\beta U^I[H_j(x, y:s)f^j(s)]V_{2n-I+1}[v^\gamma(x:s)]\, ds. \end{aligned}
$$

[7] For an argument such as this cf. W. W. Elliott, loc. cit.

[8] M. E. Bounitzky, *Sur la fonction de Green des équations différentielles linéaires ordinaires*, Journal de Mathématiques, (6), vol. 5(1909), pp. 65–125.

By (5.4), since the $f^j(t)$ are arbitrary, we have

$$\int_\alpha^\beta V_{2n-i+1}[v^\gamma(x:s)][-A_c^i(s,t)H_j^c(a,y:t) - B_c^i(s,t)H_j^c(b,y:t)]\,ds \tag{6.15}$$
$$= v_c^\gamma(a:t)H_j^c(a,y:t) - v_c^\gamma(b:t)H_j^c(b,y:t).$$

(6.14) then becomes, if we apply (5.1), (6.3), and the first equation of (6.4),

$$v_j^\gamma(y:t) + \left[\int_a^b\int_\alpha^\beta v_i^\gamma(x:s)\psi_\delta^i(x:s)\,dx\,ds - \int_\alpha^\beta V_i^\gamma(s)\rho_\delta^i(s)\,ds\right]v_j^\delta(y:t) = 0.$$

Since the $v_i^\gamma(x:s)$ are linearly independent, we obtain (6.6) as a condition on the $\psi_\delta^i(x:s)$ and the $\rho_\delta^i(s)$ for (6.5) to have a solution. This condition can be satisfied by a variety of choices of these functions; for example, the $\psi_\delta^i(x:s)$ as the negative of sets bi-orthogonal to the $v_i^\delta(x:s)$, and the $\rho_\delta^i(s)$ zero.

The continuity of the $G_j^i(x,y:s,t)$ is not apparent from equations (6.5), but if the first of that set is solved by means of (2.1), we find after simplification that

$$G_j^i(x,y:s,t) = R_c^i(x,y:s)G_j^c(y,y:s,t) + \int_\alpha^\beta S_c^i(x,y:s,r)G_j^c(y,y:r,t)\,dr$$
$$+ S_c^i(x,y:s,t)\begin{Bmatrix} H_j^c(y^+,y:t),\ x > y \\ H_j^c(y^-,y:t),\ x < y\end{Bmatrix} + \omega_\gamma^i(x:s)v_j^\gamma(y:t)$$

in which the $H_j^c(y^+,y:t)$ and $H_j^c(y^-,y:t)$ are the multipliers of $R_c^i(x,y:s)$ in (6.13) in the indicated regions. Since all other functions in the above are continuous, and since $S_j^i(x,y:s,t)$ vanishes at $x = y$, the $G_j^i(x,y:s,t)$ are continuous at $x = y$ and thus throughout I_{xyst}.

The $G_j^i(x,y:s,t)$ are not unique; in fact, by an analogous proof to the corresponding theorem in Elliott's paper the following theorem can be proved.

THEOREM 6.2. *If $G_{1j}^i(x,y:s,t)$ is any set of generalized Green's functions, then any other set $G_j^i(x,y:s,t)$ may be written in the form*

$$G_j^i(x,y:s,t) = G_{1j}^i(x,y:s,t) + u_\delta^i(x:s)c_j^\delta(y:t) + d_\delta^i(x:s)v_j^\delta(y:t) \tag{6.16}$$

for some choice of $c_j^\delta(y:t)$ and $d_\delta^i(x:s)$.

The following theorem is of interest in showing the connection that exists between a system of the type we are considering here and its adjoint.

THEOREM 6.3. *The sets of functions $H_j^i(y,x:s)$ and $G_j^i(y,x:t,s)$ are generalized Green's functions for the adjoint system* (5.7).

To show this let $u^i(x:s)$ in Green's formula (5.5) be the $G_j^i(x,y:s,t)$ satisfying (6.5), and the $v_i(x:s)$ be any solution of

$$M_i[v(x:s)] = \bar{f}^i(x:s), \qquad V_i[v(x:s)] = 0, \tag{6.17}$$

the $\bar{f}_i(x:s)$ being such that the conditions for a solution are satisfied. After simplification involving (6.3), (6.4), (6.5), (6.15), and (6.17) we have

$$v_j(y:t) = \int_a^b H_j^c(x, y:t) f_c(x:t)\,dx + \int_a^b \int_\alpha^\beta G_j^c(x, y:s, t)\bar{f}_c(x:s)\,dx\,ds + B_\delta v_j^\delta(y:t),$$

in which

$$B_\delta = \int_\alpha^\beta V_{2n-j+1}[v(x:s)]\rho_\delta^j(s)\,ds + \int_a^b \int_\alpha^\beta v_j(x:s)\psi_\delta^j(x:s)\,dx\,ds.$$

Thus the definition (6.1) of generalized Green's functions is satisfied and the theorem is proved.

7. **The auxiliary functions.** We define the auxiliary sets of continuous functions $P_j^i(x:s)$ and $Q_j^i(x:s, t)$ to be any two sets of functions, independent of the $X^i(s)$ in (5.10), which will enable us to write the solution of that system in the form

$$u^i(x:s) = P_j^i(x:s)X^j(s) + \int_\alpha^\beta Q_j^i(x:s, t)X^j(t)\,dt + B^\delta u_\delta^i(x:s), \tag{7.1}$$

whenever the $X^i(s)$ satisfy (5.12), the B^δ being arbitrary constants.

Such a set $X^i(s)$ is given by

$$X^i(s) = \begin{vmatrix} g^i(s) & V_i^1(s) & \cdots & V_i^k(s) \\ g^c V_c^1 & V_c^1 V_c^1 & \cdots & V_c^1 V_c^k \\ \cdots & \cdots & \cdots & \cdots \\ g^c V_c^k & V_c^k V_c^1 & \cdots & V_c^k V_c^k \end{vmatrix} \tag{7.2}$$

in which the $g^i(s)$ are arbitrary continuous functions in J_s and

$$g^c V_c^\delta = \int_\alpha^\beta g^c(s) V_c^\delta(s)\,ds,$$

$$V_c^\gamma V_c^\delta = \int_\alpha^\beta V_c^\gamma(s) V_c^\delta(s)\,ds.$$

These last quantities are unique and form linearly independent sets by virtue of Theorems 5.3 and 5.4. Thus the principal minor in (7.2) does not vanish. If these values of $X^i(s)$ are substituted in (7.1), and if the procedure used in Theorem 6.1 is followed, the following theorem is obtained.

THEOREM 7.1. *There exist sets of auxiliary functions $P_j^i(x:s)$ and $Q_j^i(x:s, t)$ which satisfy*

$$\begin{gathered} \frac{\partial}{\partial x} P_j^i(x:s) + \Phi_c^i(x:s)P_j^c(x:s) = 0, \\ \alpha_c^i(s)P_j^c(a:s) + \beta_c^i(s)P_j^c(b:s) = \delta_j^i \end{gathered} \tag{7.3}$$

and

$$(7.4)\quad \begin{aligned} L^i_{xs}[Q_j(x:s,\,t)] &= -K^i_c(x:s,\,t)P^c_j(x:t) + a^i_\gamma(x:s)V^\gamma_j(t), \\ U^i_{xs}[Q_j(x:s,\,t)] &= -A^i_c(s,\,t)P^c_j(a:t) - B^i_c(s,\,t)P^c_j(b:t) + b^i_\gamma(s)V^\gamma_j(t) \end{aligned}$$

the $a^i_\gamma(x:s)$ and the $b^i_\gamma(s)$ being subject to the conditions

$$(7.5)\quad \int_a^b\int_\alpha^\beta v^\gamma_i(x:s)a^i_\delta(x:s)\,dx\,ds - \int_\alpha^\beta V^\gamma_i(s)b^i_\delta(s)\,ds = \delta^\gamma_\delta.$$

A set of functions satisfying (7.3) is given by

$$(7.6)\quad P^i_j(x:s) = R^i_c(x,\,y:s)\Gamma^c_j(y:s),$$

and these functions are unique since the differential system (3.4) is incompatible and since $R^i_c(x,\,y:s)\Gamma^c_j(y:s)$ may be shown independent of y by differentiation with respect to y.

The set $Q^i_j(x:s,\,t)$ is not unique. It may be shown by a procedure entirely similar to that needed to prove Theorem 6.2 that the following theorem is true.

THEOREM 7.2. *If $Q^i_{1j}(x:s,\,t)$ is one set of auxiliary functions, any other set may be written*

$$(7.7)\quad Q^i_j(x:s,\,t) = Q^i_{1j}(x:s,\,t) + u^i_\delta(x:s)c^\delta_j(t) + d^i_\delta(x:s)V^\delta_j(t),$$

for some choice of $c^\delta_j(t)$ and $d^i_\delta(x:s)$.

If we solve the system (7.4) by the procedure of §§2 and 3, the following theorem can be proved.

THEOREM 7.3. *The $Q^i_j(x:s,\,t)$ are determinable in the form*

$$(7.8)\quad \begin{aligned} Q^i_j(x:s,\,t) = \Big[R^i_c(x,\,y:s)q^c_d(y:s,\,t) + \int_\alpha^\beta S^i_c(x,\,y:s,\,r)q^c_d(y:r,\,t)\,dr\Big]\Gamma^d_j(y:t) \\ + S^i_c(x,\,y:s,\,t)\Gamma^c_j(y:t) + \omega^i_\delta(x,\,y:s)V^\delta_j(t), \end{aligned}$$

where the $\omega^i_\delta(x,\,y:s)$ are defined by (2.17) for $f^i(x:s) = a^i_\delta(x:s)$, and the $q^i_j(y:s,\,t)$ are that pseudo-resolvent[9] *system to the kernel system $P^i_j(y:s,\,t)$ of (3.5) which satisfies*

$$(7.9)\quad \begin{aligned} q^i_j(y:s,\,t) = P^i_j(y:s,\,t) + \Gamma^i_c(y:s)\{b^c_\delta(s) - U^c[\omega_\delta(x,\,y:s)]\}V^\delta_d(t)g^d_j(y:t) \\ + \int_\alpha^\beta P^i_c(y:s,\,r)q^c_j(y:r,\,t)\,dr, \end{aligned}$$

and

$$(7.10)\quad q^i_j(y:s,\,t) = P^i_j(y:s,\,t) - u^i_\delta(y:s)\Psi^\delta_j(y:t) + \int_\alpha^\beta q^i_c(y:s,\,r)P^c_j(y:r,\,t)\,dr,$$

[9] W. A. Hurwitz, loc. cit.

the $\Psi_j^\delta(y:t)$ *being given by*

$$\Psi_j^\delta(y:t) = -\int_\alpha^\beta u_\delta^c(y:r)q_j^c(y:r,t)\,dr + \int_\alpha^\beta u_\delta^c(y:r)P_j^c(y:r,t)\,dr + \int_\alpha^\beta\int_\alpha^\beta u_\delta^c(y:r)q_d^c(y:r,s)P_j^d(y:s,t)\,dr\,ds, \tag{7.11}$$

provided that the $u_\delta^i(y:s)$ *form a normalized orthogonal set.*

To show this it is only necessary to solve (7.4) by (2.1), simplify and apply the boundary conditions. If the result is multiplied by $\Gamma_i^c(y:s)$ and $g_d^j(y:t)$, and if we let

$$Q_c^i(y:s,t)g_j^c(y:t) = q_j^i(y:s,t), \tag{7.12}$$

we obtain (7.9). Green's formula enables us to show that

$$\int_\alpha^\beta -\Gamma_c^i(y:s)\{b_\delta^c(s) - U^c[\omega_\delta(x,y:s)]\}V_d^\gamma(s)g_i^d(y:s)\,ds = \delta_\delta^\gamma. \tag{7.13}$$

(7.10) and (7.11) may then be derived from (7.9) by a procedure like that used by W. A. Hurwitz[10] in obtaining the conditions under which a function $l(x, y)$ will be a pseudo-resolvent to the kernel of a Fredholm integral equation. The $\Psi_j^\delta(y:t)$ satisfy a relation similar to (7.13) with the $u_\delta^j(y:t)$ if these latter functions form a normalized orthogonal set. If then (7.12) is solved for $Q_c^i(y:s,t)$, we have the desired result.

The generalized Green's functions defined by (6.4) and (6.5) and the auxiliary functions defined by (7.3) and (7.4) cannot be used together to write the solution of (1.3), when it exists, in the form (1.5) unless certain simple restrictions are imposed on the $a_\delta^i(x:s)$ and $b_\delta^i(s)$ of (7.4), if we suppose the $\psi_\delta^i(x:s)$ and $\rho_\delta^i(s)$ of (6.5) as given.

THEOREM 7.4. *If* (5.8) *is satisfied, the solution of* (1.3) *can be written in the form* (1.5) *if and only if the* $a_\delta^i(x:s)$ *and* $b_\delta^i(s)$ *in* (7.4) *are given by*

$$a_\delta^i(x:s) = -\psi_\delta^i(x:s); \qquad b_\delta^i(s) = -\rho_\delta^i(s), \tag{7.14}$$

the $\psi_\delta^i(x:s)$ *and* $\rho_\delta^i(s)$ *satisfying* (6.5).

That these conditions are sufficient is apparent from (1.5) if we apply the L^i and U^i and consider (5.8). To show them necessary consider the sets of functions

$$f^i(x:s) = -\frac{\begin{vmatrix} 0 & v_i^1(x:s) & \cdots & v_i^k(x:s) \\ c^1 & v_a^1 v_a^1 & \cdots & v_a^1 v_a^k \\ \cdots & \cdots & \cdots & \cdots \\ c^k & v_a^k v_a^1 & \cdots & v_a^k v_a^k \end{vmatrix}}{|v_a^\delta v_a^\gamma|},$$

[10] W. A. Hurwitz, loc. cit.

and

$$X^i(s) = \frac{\begin{vmatrix} 0 & V_i^1(s) & \cdots & V_i^k(s) \\ c^1 & V_i^1 V_i^1 & \cdots & V_i^1 V_i^k \\ \cdots & \cdots & \cdots & \cdots \\ c^k & V_i^k V_i^1 & \cdots & V_i^k V_i^k \end{vmatrix}}{| V_i^\delta V_i^\gamma |},$$

in which the c^δ are arbitrary constants. Then

$$(7.15)\qquad \begin{aligned} \int_a^b \int_\alpha^\beta f^i(x:s) v_i^\delta(x:s)\, dx\, ds &= c^\delta, \\ \int_\alpha^\beta V_i^\delta(s) X^i(s)\, ds &= c^\delta, \end{aligned}$$

and (5.8) is satisfied. With these values in (1.3), if the L^i are applied, we find that

$$(7.16)\qquad \begin{aligned} \int_a^b A_j^i(x, y:s) f^j(y:s)\, dy + \int_a^b \int_\alpha^\beta B_j^i(x, y:s, t) f^j(y:t)\, dy\, dt \\ + C_j^i(x:s) X^j(s) + \int_\alpha^\beta D_j^i(x:s, t) X^j(t)\, dt = 0. \end{aligned}$$

In the proofs of Theorems 6.1 and 7.1 we have shown that necessarily

$$A_j^i(x, y:s) = 0, \qquad B_j^i(x, y:s, t) = \psi_\delta^i(x:s) v_j^\delta(y:t),$$
$$C_j^i(x:s) = 0, \qquad D_j^i(x:s, t) = a_\delta^i(x:s) V_j^\delta(t).$$

The equation (7.16) is thus reduced with the aid of (7.15) to

$$c^\delta[\psi_\delta^i(x:s) + a_\delta^i(x:s)] = 0$$

which equation, since the c^δ are arbitrary, gives the first of the above conditions. The second of these conditions follows from a similar argument if we apply the U^i to (1.5). We note in addition that if (6.6) is satisfied, these values of $a_\delta^i(x:s)$ and $b_\delta^i(s)$, given by (7.14), will satisfy (7.5).

Duke University and North Carolina State College.

THE RING OF AUTOMORPHISMS OF AN ABELIAN GROUP

By Max Shiffman

Let A be an Abelian group (written additively) all of whose elements have a finite order. For the purposes of this study, it suffices to suppose that A is a primary group, of characteristic p. We shall consider only a restricted class of primary groups, namely, those for which every element different from 0 has a finite height.[1]

A first step in the study of the automorphisms is to describe those subgroups N which are mapped into themselves by all the (proper and improper) automorphisms α of A, $N\alpha \subset N$. These will be called *normal* subgroups.[2] The construction of all the normal subgroups of A is contained in Part I of this paper.[3] Every normal subgroup turns out to be generated by, and to be the intersection of, irreducible normal subgroups (Theorems 4, 5).

Let $\mathfrak{o}$ be the ring of all the automorphisms of A. In Part II, we establish a one-to-one correspondence between the normal subgroups of A and certain two-sided ideals of $\mathfrak{o}$, the *normal* ideals. A right normal ideal $\mathfrak{r}$ is a largest ideal which annihilates a given ideal $\mathfrak{a}$ on the right, i.e., for which $\mathfrak{a} \cdot \mathfrak{r} = 0$. To every normal subgroup N of A corresponds the totality of automorphisms in $\mathfrak{o}$ which map N into 0. This is shown to be a right normal ideal, and an inverse correspondence is established. Similarly for left normal ideals. Theorems 6, 7, 8 and 9 form the main results.

The correspondence established between normal subgroups of A and normal ideals of $\mathfrak{o}$ permits carrying over to normal ideals theorems on normal subgroups (Theorems 10–14). The most noteworthy of these results are:

the join and intersection of normal ideals are normal ideals;

the two distributive laws hold for normal ideals; and

every normal ideal is the join and also intersection of irreducible normal ideals.

A similar theory might be developed for the group $\mathfrak{G}$ of proper automorphisms of A, but it suffers from various defects. It is pointed out in §8 that such a

Received October 6, 1939.

[1] The height of an element a is the largest number m for which the equation $a = p^m x$ has a solution x in A.

[2] The names "regular characteristic", "$\mathfrak{o}$-characteristic", etc. have been used for these subgroups.

[3] Characteristic subgroups have been described by Miller, Shoda, and Baer. Cf. G. A. Miller, *Determination of all the characteristic subgroups of an Abelian group*, Quart. Journ. of Math., vol. 50(1923), pp. 54–62; K. Shoda, *Über die charakteristischen Untergruppen einer endlichen Abelschen Gruppe*, Math. Zeitsch., vol. 31(1930), pp. 611–624; R. Baer, *Types of elements and the characteristic subgroups of Abelian groups*, Proc. London Math. Soc., (2), vol. 39(1935), pp. 481–514. Our procedure is closely akin to that of Baer.

theory can only be understood if the group $\mathfrak{G}$ is considered as imbedded in the ring $\mathfrak{o}$.

It is instructive to compare the theory developed herein with the Galois theory of fields. In the latter, a one-to-one correspondence is established between normal subfields of a given field and invariant subgroups of the group of automorphisms of the field. In the theory of this paper, normal subgroups are mapped into normal ideals of $\mathfrak{o}$; i.e., only special ideals of $\mathfrak{o}$ are covered, whereas in the Galois theory every invariant subgroup is covered. Also, the Galois theory uses the structure of the group of automorphisms to determine that of the original field, whereas our theory should be applied to study the ring $\mathfrak{o}$, use of the knowledge of the Abelian group A being made.

Elements of A will be denoted by italic letters $a, b, c, \cdots$, automorphisms of A by Greek letters $\alpha, \beta, \gamma, \cdots$, and ideals in $\mathfrak{o}$ by German letters $\mathfrak{a}, \mathfrak{b}, \mathfrak{l}, \mathfrak{r}, \cdots$. The effect of applying the automorphism α on the element a is written $a\alpha$. The symbol $\{B, C, \cdots\}$ will designate the group or ideal generated by the sets of elements $B, C, \cdots$; $[B, C, \cdots]$ the intersection of the sets $B, C, \cdots$; $B \subset C$ or $C \supset B$ means that every element in B belongs to C.

Part I. Normal subgroups

1. **Introduction.** We have restricted the Abelian group A to be a primary Abelian group all of whose elements (different from 0) have a finite height. With each element a ($\neq 0$) of A there are therefore associated two numbers: $n(a)$, its exponent *order*, such that $p^{n(a)}a = 0$ but $p^{n(a)-1}a \neq 0$; $h(a)$, its exponent *height*, such that $a \subset p^{h(a)}A$ but $a \not\subset p^{h(a)+1}A$. These will hereafter be called simply the order and height respectively of a. The element 0 will be considered as having the order 0 and the height ∞.

We shall now set down, without proof, some well-known basic theorems for Abelian groups of this type.[4] Let A_n denote the subgroup of A consisting of all the elements x of A whose order is not larger than n, $n(x) \leqq n$; the subgroup $p^h A_k$, where $h \leqq k$, consists of all those elements x of A for which $h(x) \geqq h$, $n(x) \leqq k - h$. Consider the subgroup A_1. Since the order of every element ($\neq 0$) of A_1 is 1, A_1 may be written as a direct sum of cyclic groups of order 1. Determine this direct sum in the following way: let $c_{g\nu}$ $(g = 1, \cdots)$ be any basis of $p^{\nu-1}A_\nu / p^\nu A_{\nu+1}$; then set $A_1 = \sum_{\nu=1}^{\infty} \sum_g \{c_{g\nu}\} + A_1'$. Since $h(c_{g\nu}) = \nu - 1$, determine any element $b_{g\nu}$ of A such that $c_{g\nu} = p^{\nu-1}b_{g\nu}$. It is a basic result that the elements $b_{g\nu}$ of A are independent, so that one may form $\sum_{\nu=1}^{m-1} \sum_g \{b_{g\nu}\}$ and $\sum_{\nu=m}^{\infty} \sum_g \{b_{g\nu}\}$. Furthermore, let B^m be the "closure" of $\sum_{\nu=m}^{\infty} \sum_g \{b_{g\nu}\}$, i.e., the totality of elements x of A such that, for each i, there is an element x_i of

[4] Cf. H. Prüfer, *Zerlegbarkeit der abzählbaren primären Abelschen Gruppen*, Math. Zeitsch., vol. 17(1923), pp. 35–61; and §1 of Baer, *Primary Abelian groups and their automorphisms*, Amer. Jour. Math., vol. 59(1937), pp. 99–117.

$\sum_{\nu=m}^{\infty} \sum_{g} \{b_{g\nu}\}$ for which $h(x - x_i) \geqq i$. Then it is a fundamental theorem that A admits the following direct decomposition, for each m,[5]

$$A = \sum_{\nu=1}^{m-1} \sum_{g} \{b_{g\nu}\} + B^m.$$

Also, if x is any element of B^m, then $h(x) + n(x) \geqq m$.

The elements $b_{g\nu}$ have the property that $n(b_{g\nu}) = \nu$ and $h(p^{\nu-1}b_{g\nu}) = \nu - 1$. Elements with this property will be called primitive elements. That is, the element b of A is a *primitive* element if $h(p^i b) = i$ for each $i = 0, 1, \cdots, n(b) - 1$ (it suffices to know this only for $i = n(b) - 1$). Because of the freedom in the choice of $b_{g\nu}$ in the direct decomposition above, any primitive element of order ν may be selected as a $b_{g\nu}$. In fact, if $b_1, \cdots, b_t$ are primitive elements of different orders $n_1, \cdots, n_t$, then $p^{n_1-1}b_1, \cdots, p^{n_t-1}b_t$ can be chosen as basis elements of $p^{n_1-1}A_{n_1}/p^{n_1}A_{n_1+1}, \cdots, p^{n_t-1}A_{n_t}/p^{n_t}A_{n_t+1}$, respectively. Hence $b_1, \cdots, b_t$ may be chosen as elements of the type $b_{g\nu}$ in the direct decomposition above, where m is chosen larger than all the n_i. Letting A' be the subgroup generated by all the remaining basic groups in that direct decomposition, we get

$$A = \{b_1\} + \cdots + \{b_t\} + A'.$$

Thus, primitive elements of different orders form part of a basis of A.

2. **The invariants of an element. Normal subgroups.** The element a' is said to be *homotopic* to a if there is an automorphism α, proper or improper, of A which maps a into a', $a\alpha = a'$. If there is a proper automorphism β of A which takes a into a', then a' is said to be *isotopic* to a. The question of determining when two elements are isotopic or homotopic will be answered by use of the following important lemma of Baer's.[6]

LEMMA 1. *Any element a of A can be expressed in the normal form*

$$a = \sum_{i=1}^{t} p^{h_i} b_i,$$

where $b_1, \cdots, b_t$ are primitive elements with orders $n_1 + h_1, \cdots, n_t + h_t$ respectively, and $0 \leqq h_1 < \cdots < h_t$, $0 < n_1 < \cdots < n_t$.

Proof. The lemma is true if a has the order 1, for then a is a multiple of a primitive element. Suppose it has been proved for all elements of order $\leqq n$, and let $n(a) = n + 1$. Among all the multiples $p^i a$ $(i = 0, 1, \cdots, n)$, let $p^{n'} a$ be the first which is a multiple of a primitive element, $p^{n'} a = p^{n'+h'} b$. Then the element $a' = a - p^{h'} b$ has an order $= n' \leqq n$, and the lemma is

[5] A simple proof can be given using induction on the order of elements of A.

[6] See Baer, loc. cit. (footnote 3).

applicable to it. By writing $a = a' + p^{h'}b$, it is easily seen that the lemma is valid.

The set of $2t$ numbers $h_1, \cdots, h_t, n_1, \cdots, n_t$ forms a complete set of invariants of the element a, and can be expressed in terms of the heights of the multiples $p^j a$ $(j = 0, 1, \cdots, n(a) - 1)$. For, $p^j a = \sum_{i=1}^{t} p^{h_i+j} b_i$ so that $h(p^j a) = h_i + j$ for that i determined by $n_{i-1} \leqq j < n_i$.

The primitive elements $b_1, \cdots, b_t$ in the normal form of a have different orders so that the theorem of §1 may be applied. It can then be easily shown that two elements a, a' are isotopic if and only if the invariants $h_1, \cdots, h_t, n_1, \cdots, n_t$ defined in Lemma 1 are the same for both elements.[6] The question of determining when an element a' is homotopic to a will be incidentally answered in the course of §3.

A subgroup N of A will be called a *normal* subgroup if it is mapped into itself by every automorphism α (proper or improper) of A, $N\alpha \subset N$. The subgroup $p^h A_k$ is clearly a normal subgroup of A since any automorphism of A never increases the order nor decreases the height of an element. The group generated by, and the intersection of, any number of normal subgroups is again a normal subgroup. All normal subgroups will turn out to be generated by the elementary normal subgroups $p^h A_k$.

For any element a of A denote the smallest normal subgroup containing a by $N(a)$. Our method for studying the normal subgroups will be first to determine the structure of $N(a)$ and then build any normal subgroup out of these.[7]

It is clear that $N(a)$ must contain all the homotopes to a. On the other hand, the set of all the homotopes to a forms a normal subgroup of A. For, this set is a subgroup: if a' and a'' are homotopic to a, i.e., $a\alpha' = a'$ and $a\alpha'' = a''$, then $a(\alpha' + \alpha'') = a' + a''$ and $a' + a''$ is homotopic to a. This subgroup is normal: if α is any automorphism of A and a' is any homotope to a, i.e., $a\alpha' = a'$, then the image of a' is $a'' = a'\alpha = (a\alpha')\alpha = a(\alpha'\alpha)$ which is homotopic to a. Since any normal subgroup N which contains a must contain $N(a)$, we see that $N(a)$ is exactly the normal subgroup of all the homotopes to a.

3. The structure of the normal subgroup $N(a)$.

THEOREM 1. *If a has the normal form*

$$a = \sum_{i=1}^{t} p^{h_i} b_i, \qquad n(b_i) = n_i + h_i,$$

[7] A similar procedure can be used to find all the characteristic subgroups of A (a characteristic subgroup is one which is mapped into itself by all the proper automorphisms of A); determine the smallest characteristic group $C(a)$ containing the element a. It is found that $C(a) = N(a)$ unless $p = 2$ and another condition is satisfied. Thus, if $p \neq 2$, all characteristic subgroups are at the same time normal subgroups. See Baer, loc. cit. (footnote 3).

then

$$N(a) = \{p^{h_1}A_{n_1+h_1}, \cdots, p^{h_t}A_{n_t+h_t}\}.$$

Proof. Since $b_1, \cdots, b_t$ are primitive elements of different orders,

$$A = \{b_1\} + \{b_2\} + \cdots + \{b_t\} + A'.$$

Let x be any element of $A_{n_j+h_j}$, so that $n(x) \leqq n_j + h_j$. The mapping $b_j\alpha = x$, $b_i\alpha = 0$ for $i \neq j$, $A'\alpha = 0$ defines an improper automorphism α of A for which $a\alpha = \sum_{i=1}^{t} p^{h_i}(b_i\alpha) = p^{h_j}x$. Thus $N(a)$ must contain every element $p^{h_j}x$ of $p^{h_j}A_{n_j+h_j}$. Since this holds for every $j = 1, \cdots, t$, we have

$$N(a) \supset \{p^{h_1}A_{n_1+h_1}, \cdots, p^{h_t}A_{n_t+h_t}\}.$$

On the other hand, the group on the right side of the equation is a normal subgroup containing a, so that

$$N(a) \subset \{p^{h_1}A_{n_1+h_1}, \cdots, p^{h_t}A_{n_t+h_t}\}.$$

The equality sign therefore holds and the theorem is proved.

The element a' is homotopic to a if $a' \subset N(a)$. By Theorem 4 below, this requires that each component $p^{h'}b'$ in the normal form of a' be contained in at least one of the subgroups $p^{h_j}A_{n_j+h_j}$ $(j = 1, \cdots, t)$.

4. **The structure of normal subgroups.** Let N be any normal subgroup of A. N certainly is generated by the normal groups $N(c)$ for all elements c in N. By Theorem 1, N is generated by groups of the type p^hA_k. For a given k, let $h_N(k)$ be the smallest h such that N contains p^hA_k; let $n_N(k)$ be the corresponding n, $n_N(k) + h_N(k) = k$. Then N is generated by the groups $p^{h_N(k)}A_k$ $(k = 1, 2, 3, \cdots)$.

From the relations $p^hA_k \subset p^hA_{k+1}$, $p^hA_k \supset p^{h+1}A_{k+1}$, it follows, by setting $h = h_N(k + 1)$ in the first and $h = h_N(k)$ in the second, that $h_N(k) \leqq h_N(k + 1) \leqq h_N(k) + 1$. Hence, we also have $1 + n_N(k) \geqq n_N(k + 1) \geqq n_N(k)$.

Suppose that no primitive elements of order k' exist in A. Then $p^{h'}A_{k'}$ is generated by all the groups p^hA_k $(k = n + h)$ for which $h \geqq h'$, $n \leqq n'$ $(= k' - h')$, and such that primitive elements of order k exist in A. For, on the one hand, any element a of $p^{h'}A_{k'}$ has the normal form $\sum_{i=1}^{t} p^{h_i}b_i$, $n(b_i) = n_i + h_i$, where $h_i \geqq h_1 \geqq h'$, $n_i \leqq n_t \leqq n'$; and on the other hand, $p^{h'}A_{k'} \supset p^hA_{n+h}$ for such h's and n's. Thus, N is generated by $p^{h_N(k)}A_k$ for all relevant k's. By a relevant order k is meant one such that primitive elements of order k exist in A.[8] Hereafter we consider only such k's.

The above inequalities for $h_N(k)$, $n_N(k)$ are equivalent to $h_N(k_1) \leqq h_N(k_2)$,

[8] Let r be the number of independent primitive elements of order k in A. The order of $p^{k-1}A_k/p^kA_{k+1}$ is p^r or r according as r is finite or infinite respectively.

$n_N(k_1) \leqq n_N(k_2)$ for $k_1 < k_2$. The numbers $h_N(k)$, $n_N(k)$ form a complete set of invariants of N.

THEOREM 2.[9] *Let $h(k)$, $n(k)$ be a set of non-negative numbers defined for all relevant k's. The necessary and sufficient condition that there exist a normal subgroup N of A for which $h_N(k) = h(k)$, $n_N(k) = n(k)$ is that*
(1) $h(k) + n(k) = k$;
(2) $h(k_1) \leqq h(k_2)$, $n(k_1) \leqq n(k_2)$ *for* $k_1 < k_2$.
Then N is unique and is the group generated by all the groups $p^{h(k)}A_k$,

$$N = \{\cdots, p^{h(k)}A_k, \cdots\}.$$

Proof. The necessity has already been proved, so that only the sufficiency remains. Define N as the normal subgroup generated by all the groups $p^{h(k)}A_k$. Then $h_N(k) \leqq h(k)$, $n_N(k) \geqq n(k)$ and we must show that the equality sign holds. Let b be a primitive element of order k, and let h be a number $< h(k)$, so that $n = k - h > k - h(k) = n(k)$. The element $p^h b$ can belong to N only if

$$p^h b = \sum_{j \leqq k} p^{h(j)} a_j + \sum_{j > k} p^{h(j)} a_j,$$

where $n(a_j) \leqq j$; by multiplying by p^{n-1} this requires that

$$p^{k-1} b = \sum_{j \leqq k} p^{n+h(j)-1} a_j + \sum_{j > k} p^{n+h(j)-1} a_j.$$

Since $n > n(k) \geqq n(j)$ for $j \leqq k$ by condition (2), it follows that

$$p^{k-1} b = \sum_{j > k} p^{n+h(j)-1} a_j.$$

By condition (2) again, $h(j) \geqq h(k) > h$ for $j > k$, and the height of the right side is larger than $n + h - 1 = k - 1$. But this is contrary to the primitivity of b. Thus, $p^h b$ and therefore $p^h A_k$ does not belong to N if $h < h(k)$. Hence, $h_N(k) = h(k)$, $n_N(k) = n(k)$, and the sufficiency is proved.

The uniqueness is a consequence of the representation of every normal subgroup N as generated by all the $p^{h_N(k)}A_k$. This completes the proof of the theorem.

COROLLARY 1. *If A is a direct sum of cyclic groups, $A = \sum_\nu \sum_\sigma \{b_{\sigma\nu}\}$, where $b_{\sigma\nu}$ has the order ν, then*

$$N = \sum_\nu \sum_\sigma \{p^{h_N(\nu)} b_{\sigma\nu}\}.$$

Proof. This follows by inserting the expression

$$p^{h_N(k)} A_k = \sum_{\nu \leqq k} \sum_\sigma \{p^{h_N(k)} b_{\sigma\nu}\} + \sum_{\nu > k} \sum_\sigma \{p^{\nu - n_N(k)} b_{\sigma\nu}\}$$

[9] Cf. Baer, loc. cit. (footnote 3).

in $N = \{ \cdots, p^{h_N(k)}A_k, \cdots \}$ and obtaining the minimum coefficient of a given $b_{\sigma\nu}$.

THEOREM 3. *Let* $\cdots, N_i, \cdots$ *be a set of normal subgroups, and let* $N = \{ \cdots, N_i, \cdots \}$, $M = [\cdots, N_i, \cdots]$. *Then*

(a) $h_N(k) = \min_i h_{N_i}(k)$, $n_N(k) = \max_i n_{N_i}(k)$;

(b) $h_M(k) = \max_i h_{N_i}(k)$, $n_M(k) = \min_i n_{N_i}(k)$.

Proof. To prove (a) let $h(k) = \min_i h_{N_i}(k)$, $n(k) = \max_i n_{N_i}(k)$. They are non-negative numbers and satisfy condition (1) of Theorem 2 since $n_{N_i}(k) = k - h_{N_i}(k)$ for every N_i. To verify condition (2) of Theorem 2, let $k_1 < k_2$, and determine $N_{j'}$, $N_{j''}$ so that

$$n_{N_{j'}}(k_1) = \max_i n_{N_i}(k_1), \qquad h_{N_{j''}}(k_2) = \min_i h_{N_i}(k_2).$$

Then

$$h(k_1) = \min_i h_{N_i}(k_1) \leqq h_{N_{j''}}(k_1) \leqq h_{N_{j''}}(k_2) = h(k_2),$$

and

$$n(k_1) = n_{N_{j'}}(k_1) \leqq n_{N_{j'}}(k_2) \leqq \max_i n_{N_i}(k_2) = n(k_2).$$

There exists, therefore, a normal subgroup N' with $h_{N'}(k) = h(k)$, $n_{N'}(k) = n(k)$.

Since $h_{N'}(k) \leqq h_{N_i}(k)$ for every k, $N' \supset N_i$; varying i, we obtain $N' \supset \{ \cdots, N_i, \cdots \} = N$. On the other hand, $h_N(k) \leqq h_{N_i}(k)$ because $N \supset N_i$; varying i, we obtain $h_N(k) \leqq \min_i h_{N_i}(k) = h_{N'}(k)$ for each k, or $N \supset N'$. Thus, $N = N'$ and (a) is proved.

Next consider (b). Set $h(k) = \max_i h_{N_i}(k)$, $n(k) = \min_i n_{N_i}(k)$. As in the proof of (a), but reversing the rôles of h and n, there is a normal subgroup M' with $h_{M'}(k) = h(k)$, $n_{M'}(k) = n(k)$. Since $h_{M'}(k) \geqq h_{N_i}(k)$ for every k, $M' \subset N_i$; varying i, we obtain $M' \subset [\cdots, N_i, \cdots] = M$. On the other hand, $h_M(k) \geqq h_{N_i}(k)$ because $M \subset N_i$; varying i, we get $h_M(k) \geqq \max_i h_{N_i}(k) = h_{M'}(k)$ for each k, so that $M \subset M'$. Thus $M = M'$ and (b) is proved.

COROLLARY 2. *If* $p^h A_k = \{ \cdots, N_i, \cdots \}$, *then at least one of the normal subgroups* N_i *is* $p^h A_k$.

Proof. By the previous theorem, $h = \min_i h_{N_i}(k) = h_{N_j}(k)$ for some N_j. Hence, $N_j \supset p^h A_k$. But $p^h A_k \supset N_j$, so that finally $N_j = p^h A_k$.

COROLLARY 3. *The following distributive laws hold:*[10]

$$[\{N_1, N_2\}, N_3] = \{[N_1, N_3], [N_2, N_3]\}, \quad \{[N_1, N_2], N_3\} = [\{N_1, N_3\}, \{N_2, N_3\}].$$

[10] The extended distributive laws with any number of terms also hold.

Proof. By the previous theorem, we must prove that max [min $(h_1, h_2), h_3$] = min [max (h_1, h_3), max (h_2, h_3)] and that a similar equation obtained by interchanging max and min holds. This is easily done.

The normal subgroup N has been represented as the group generated by all the $p^{h_N(k)}A_k$, where $h_N(k_1) \leqq h_N(k_2)$, $n_N(k_1) \leqq n_N(k_2)$ for $k_1 < k_2$. We shall now obtain a minimum representation of N. Suppose that $h_N(k_1) = h_N(k_2)$. Then $p^{h_N(k_2)}A_{k_2} \supset p^{h_N(k_1)}A_{k_1}$, and $p^{h_N(k_1)}A_{k_1}$ may be deleted from the set of groups generating N. Similarly, if $n_N(k_1) = n_N(k_2)$, $p^{h_N(k_2)}A_{k_2}$ may be deleted. In carrying through these deletions, it may happen that $h_N(k)$ is constant, $= h$, for all $k \geqq s$. In that case, all the groups $p^{h_N(k)}A_k$ for $k \geqq s$ are to be deleted and replaced by p^hA. This will be written p^hA_∞; n will be taken equal to ∞, and ∞ will be called a relevant order of A (∞ is a relevant order, then, if the elements of A have unbounded orders). We thus obtain N as generated by the finite or infinite set of remaining groups $p^{h_i}A_{n_i+h_i}$, where

$$0 \leqq h_1 < \cdots < h_i < \cdots, \qquad 0 < n_1 < \cdots < n_i < \cdots \leqq \infty,$$

and $n_i + h_i = k_i$ are relevant orders. Hence, we have

THEOREM 4. *Every normal subgroup N can be represented in a unique manner as*

$$N = \{p^{h_1}A_{n_1+h_1}, \cdots, p^{h_i}A_{n_i+h_i}, \cdots\},$$

where

(1) $0 \leqq h_1 < \cdots < h_i < \cdots, 0 < n_1 < \cdots < n_i < \cdots \leqq \infty$; *and*

(2) $n_i + h_i$ *are all relevant orders.*

The group p^hA_k is contained in N if and only if it is contained in one of the generating groups $p^{h_i}A_{n_i+h_i}$. Conversely, any finite or infinite set of numbers h_i, n_i satisfying (1), (2) *uniquely determines a normal subgroup N.*

The similarity of Theorems 1 and 4 leads immediately to

COROLLARY 4. *If the invariants h_i, n_i of N are finite in number ($i = 1, 2, \cdots, t$) (i.e., if either $h_N(k)$ or $n_N(k)$ is bounded for all k), then there is an element a such that either $N = N(a)$ or $N = \{N(a), p^{h_t}A\}$ according as n_t is finite or infinite respectively.*

COROLLARY 5. *If the orders of the elements of A are bounded, every normal subgroup N is a subgroup $N(a)$ for some element a.*

If the invariants h_i, n_i are infinite in number, N may be considered as the smallest normal subgroup containing the purely formal sum $\sum_{i=1}^{\infty} p^{h_i}b_i$, where the b_i are primitive elements of order $n_i + h_i$.

In all the above, the invariants $h(k)$ and $n(k)$ have played equivalent rôles. We may therefore define the *dual* N' to any normal subgroup N by $h_{N'}(k) =$

$n_N(k)$, $n_{N'}(k) = h_N(k)$. It is evident that N is likewise dual to N'; and by Theorem 3, that

$$\{\cdots, N_i, \cdots\}' = [\cdots, N_i', \cdots] \quad \text{and} \quad [\cdots, N_i, \cdots]' = \{\cdots, N_i', \cdots\}.$$

Denote the dual of $p^n A_k$ by ${}_nA^k$; by calculating the invariants $h(l)$, $n(l)$ of $p^n A_k$, one easily obtains

$${}_nA^k = \{A_n, p^k A\}.$$

Thus, ${}_nA^k$ consists of all elements x such that $p^n x \subset p^k A$. In particular, ${}_0A^k = p^k A$ and ${}_nA^\infty = A_n$.

By taking the dual of the various theorems, new theorems are obtained. The dual of Theorem 2 yields $N' = [\cdots, {}_{h_N(k)}A^k, \cdots]$; replace N by N', so that $N = [\cdots, {}_{n_N(k)}A^k, \cdots]$. Delete as in the proof of Theorem 4, obtaining $N = [\cdots, {}_{n_i}A^{n_i+h_i}, \cdots]$, etc. These results are summarized in

THEOREM 5. *Every normal subgroup N can be represented uniquely as the intersection of all the groups ${}_{n_N(k)}A^k$,*

$$N = [\cdots, {}_{n_N(k)}A^k, \cdots].$$

We also have a unique representation in the form

$$N = [\cdots, {}_{n_i}A^{n_i+h_i}, \cdots],$$

where

(1) $0 \leqq n_1 < \cdots < n_i < \cdots, 0 < h_1 < \cdots < h_i < \cdots \leqq \infty$; *and*

(2) $n_i + h_i$ *are relevant orders for each* i.

The group ${}_nA^k$ contains N if and only if it contains one of the groups ${}_{n_i}A^{n_i+h_i}$. Here, ${}_nA^k$ represents the dual of $p^n A_k$ and is given by

$${}_nA^k = \{A_n, p^k A\}.$$

The duals of Corollaries 2, 4 yield additional theorems.

Part II. The ring of automorphisms

5. **The correspondences.** We shall establish one-to-one correspondences between the normal subgroups of A and certain two-sided ideals of the ring $\mathfrak{o}$ of automorphisms of A. Ideals of $\mathfrak{o}$ will be denoted by $\mathfrak{i}$, $\mathfrak{j}$, $\cdots$, $\mathfrak{a}$, $\mathfrak{b}$, $\cdots$, normal subgroups of A by $N, N_1, \cdots$.

Let N be a given normal subgroup of A. Define $\mathfrak{r}(N)$ as the set of all automorphisms α of A which map each element of N into 0, and $\mathfrak{l}(N)$ as the set of all automorphisms β of A which map each element of A into an element of N:[11]

[11] Similar correspondences have been discussed by Shoda and by Baer. See Shoda, *Über die Automorphismen einer endlichen Abelschen Gruppe*, Math. Ann., vol. 100(1928), pp. 674–686; Baer, loc. cit. (footnote 4).

$\mathfrak{r}(N)$ = set of all automorphisms α such that $N\alpha = 0$;
$\mathfrak{l}(N)$ = set of all automorphisms β such that $A\beta \subset N$.
It is easily seen that both $\mathfrak{r}(N)$ and $\mathfrak{l}(N)$ are two-sided ideals.

The inverse operations to $\mathfrak{r}(\ )$ and $\mathfrak{l}(\ )$ are the following. Let $\mathfrak{i}$ be a given two-sided ideal of $\mathfrak{o}$. Define $R(\mathfrak{i})$ as the set of all elements a of A which are mapped into 0 by all the automorphisms in $\mathfrak{i}$, and $L(\mathfrak{i})$ as the group generated by all the elements of the form $a'\alpha'$, where a' is any element of A and α' any automorphism in $\mathfrak{i}$:
$R(\mathfrak{i})$ = set of all elements a such that $a\mathfrak{i} = 0$;
$L(\mathfrak{i})$ = group generated by all the elements $a' \subset A\mathfrak{i}$.
It is easily seen that both $R(\mathfrak{i})$ and $L(\mathfrak{i})$ are normal subgroups of A.

We shall try to determine when $\mathfrak{r}$ and R, $\mathfrak{l}$ and L are inverse to one another. The following Lemmas 2, 3 are immediate consequences of the definitions.

Lemma 2. (a) $R(\mathfrak{r}(N)) \supset N$, $\mathfrak{r}(R(\mathfrak{i})) \supset \mathfrak{i}$;
(b) $L(\mathfrak{l}(N)) \subset N$, $\mathfrak{l}(L(\mathfrak{i})) \supset \mathfrak{i}$.

Lemma 3. (a) *If* $N_1 \subset N_2$, *then* $\mathfrak{r}(N_1) \supset \mathfrak{r}(N_2)$, $\mathfrak{l}(N_1) \subset \mathfrak{l}(N_2)$;
(b) *if* $\mathfrak{i} \subset \mathfrak{j}$, *then* $R(\mathfrak{i}) \supset R(\mathfrak{j})$, $L(\mathfrak{i}) \subset L(\mathfrak{j})$.

Lemma 4. (a) $\mathfrak{r}(R(\mathfrak{r}(N))) = \mathfrak{r}(N)$, $R(\mathfrak{r}(R(\mathfrak{i}))) = R(\mathfrak{i})$;
(b) $\mathfrak{l}(L(\mathfrak{l}(N))) = \mathfrak{l}(N)$, $L(\mathfrak{l}(L(\mathfrak{i}))) = L(\mathfrak{i})$.

Proof. First prove (a). Apply the operator $\mathfrak{r}$ to both sides of the first inequality of Lemma 2(a). By Lemma 3(a), $\mathfrak{r}(R(\mathfrak{r}(N))) \subset \mathfrak{r}(N)$. But $\mathfrak{r}(R(\mathfrak{r}(N))) \supset \mathfrak{r}(N)$ by the second inequality of Lemma 2(a). Hence the equality sign holds. Similarly for $R(\mathfrak{r}(R(\mathfrak{i})))$.

The proof of (b) is similar to that of (a).

Thus, the operations $\mathfrak{r}$ and R are inverse to one another when applied to ideals of the type $\mathfrak{r}(N)$ or to normal subgroups of the type $R(\mathfrak{i})$. Conversely, if these operations are inverse to one another when applied to the ideal $\mathfrak{i}$, then $\mathfrak{i}$ is of the type $\mathfrak{r}(N)$. For, if $\mathfrak{r}(R(\mathfrak{i})) = \mathfrak{i}$, then $\mathfrak{i} = \mathfrak{r}(N)$ for $N = R(\mathfrak{i})$. Likewise for the other operations. The question of when the operations $\mathfrak{r}$ and R, $\mathfrak{l}$ and L are inverse to one another is therefore equivalent to the problem of determining what ideals are of the type $\mathfrak{r}(N)$ or $\mathfrak{l}(N)$ and what normal subgroups are of the type $R(\mathfrak{i})$ or $L(\mathfrak{i})$.

6. **The main theorems.** The following basic theorem shows that every normal subgroup is of the type $R(\mathfrak{i})$ and $L(\mathfrak{i})$.

Theorem 6.[12] (1) $R(\mathfrak{r}(N)) = N$.
(2) $L(\mathfrak{l}(N)) = N$.

Proof. Consider (1). By Lemma 2(a), it suffices to show that $R(\mathfrak{r}(N)) \subset N$, i.e., if $a \not\subset N$ then $a \not\subset R(\mathfrak{r}(N))$. We shall construct an automorphism α

[12] Cf. Baer, loc. cit. (footnote 4).

which belongs to $\mathfrak{r}(N)$, i.e., $N\alpha = 0$, and for which $a\alpha \neq 0$. This would show that a does not belong to $R(\mathfrak{r}(N))$.

Let $h_N(k)$, $n_N(k)$ be the invariants of N. Let $a = \sum_{i=1}^{t} p^{h_i} b_i$, $n(b_i) = n_i + h_i = k_i$, be the normal form of a. Since $a \not\subset N$, at least one $p^{h_j} b_j \not\subset N$, so that $h_j < h_N(k_j)$, $n_j > n_N(k_j)$. Now, set $A = \{b_1\} + \cdots + \{b_t\} + A' = \{b_j\} + A''$, where A'' is the direct sum of all the summands excluding $\{b_j\}$, and define the automorphism α by

$$b_j\alpha = p^{n_j - 1} b_j, \qquad A''\alpha = 0.$$

This automorphism maps a into $a\alpha = \sum_{i=1}^{t} p^{h_i} b_i \alpha = p^{h_j + n_j - 1} b_j \neq 0$. We shall show that $N\alpha = 0$.

Any element c of A has the form $c = sb_j + c''$, where $c'' \subset A''$. Let $p^h c$ be an element of $p^h A_k$, so that $n(c) \leqq k$. There are two cases to distinguish.

(a) If $k \geqq k_j$, then $p^h c = sp^h b_j + p^h c''$. Applying the automorphism α, we obtain $p^h c\alpha = sp^{h+n_j-1} b_j$ which is 0 if $h > h_j$. Thus, $p^{h_N(k)} A_k \alpha = 0$ since $h_N(k) \geqq h_N(k_j) > h_j$.

(b) If $k \leqq k_j$, then $c = sp^{k_j - k} b_j + c''$ and $p^h c = sp^{k_j - n} b_j + p^h c''$, where $n = k - h$. Applying the automorphism α, we obtain $p^h c\alpha = sp^{k_j - n + n_j - 1} b_j$ which is 0 if $n < n_j$. Thus, $p^{h_N(k)} A_k \alpha = 0$ since $n_N(k) \leqq n_N(k_j) < n_j$. Hence $N\alpha = 0$.

We have thus constructed, for any a not belonging to N, an automorphism α belonging to $\mathfrak{r}(N)$ for which $a\alpha \neq 0$. Therefore a does not belong to $R(\mathfrak{r}(N))$, and $R(\mathfrak{r}(N)) \subset N$. This proves the equality sign because of Lemma 2(a).

We now prove (2). By Lemma 2(b), it suffices to show that $L(\mathfrak{l}(N)) \supset N$, i.e., if $a \subset N$, then $a \subset L(\mathfrak{l}(N))$. Let the normal form of a be $a = \sum_{i=1}^{t} p^{h_i} b_i$, where $h_i \geqq h_N(k_i)$. Let $A = \{b_1\} + \cdots + \{b_t\} + A'$, and define the automorphism β by

$$b_i\beta = p^{h_N(k_i)} b_i \quad (i = 1, \cdots, t), \qquad A'\beta = 0.$$

Now, any element c of A has the form $c = s_1 b_1 + \cdots + s_t b_t + c'$, so that

$$c\beta = s_1 p^{h_N(k_1)} b_1 + \cdots + s_t p^{h_N(k_t)} b_t \subset N.$$

Hence $A\beta \subset N$ and β belongs to $\mathfrak{l}(N)$. The element $\sum_{i=1}^{t} p^{h_i - h_N(k_i)} b_i$ is mapped by β into $\sum_{i=1}^{t} p^{h_i - h_N(k_i)} b_i \beta = \sum_{i=1}^{t} p^{h_i} b_i = a$. Therefore $a \subset A\beta$ so that $a \subset L(\mathfrak{l}(N))$. This completes the proof of Theorem 6.

We shall now characterize those ideals of the type $\mathfrak{r}(N)$ or $\mathfrak{l}(N)$. It is clear that

$$\mathfrak{l}(N) \cdot \mathfrak{r}(N) = 0 \qquad \text{(but not } \mathfrak{r}(N) \cdot \mathfrak{l}(N) = 0!).$$

For,

$$A(\mathfrak{l}(N)\cdot\mathfrak{r}(N)) = (A\mathfrak{l}(N))\mathfrak{r}(N) \subset N\mathfrak{r}(N) = 0.$$

That is, every automorphism of $\mathfrak{l}(N)\cdot\mathfrak{r}(N)$ first maps A into N and then N into 0, and so is the 0 automorphism. This property will serve to characterize the ideals $\mathfrak{l}(N)$, $\mathfrak{r}(N)$.

LEMMA 5. *The necessary and sufficient condition that* $\mathfrak{i}\cdot\mathfrak{j} = 0$ *is that* $L(\mathfrak{i}) \subset R(\mathfrak{j})$.

Proof. If $\mathfrak{i}\cdot\mathfrak{j} = 0$, then $0 = A(\mathfrak{i}\cdot\mathfrak{j}) = (A\mathfrak{i})\mathfrak{j}$. The totality of elements $A\mathfrak{i}$, and therefore the normal subgroup $L(\mathfrak{i})$ generated by them, is mapped into 0 by all the automorphisms in $\mathfrak{j}$. That is, $L(\mathfrak{i}) \subset R(\mathfrak{j})$. Conversely, if $L(\mathfrak{i}) \subset R(\mathfrak{j})$, then $A(\mathfrak{i}\cdot\mathfrak{j}) = (A\mathfrak{i})\mathfrak{j} \subset L(\mathfrak{i})\mathfrak{j} \subset R(\mathfrak{j})\mathfrak{j} = 0$. Hence $\mathfrak{i}\cdot\mathfrak{j} = 0$.

THEOREM 7. (1) *The totality of automorphisms* α *for which* $\mathfrak{l}(N)\cdot\alpha = 0$ *is exactly* $\mathfrak{r}(N)$.

(2) *The totality of automorphisms* β *for which* $\beta\cdot\mathfrak{r}(N) = 0$ *is exactly* $\mathfrak{l}(N)$.

Proof. First consider (1). The totality of automorphisms α for which $\mathfrak{l}(N)\cdot\alpha = 0$ forms an ideal $\mathfrak{j}$ which contains $\mathfrak{r}(N)$, so that $\mathfrak{l}(N)\cdot\mathfrak{j} = 0$ and $\mathfrak{j} \supset \mathfrak{r}(N)$. By Lemma 5, $\mathfrak{l}(N)\cdot\mathfrak{j} = 0$ implies $L(\mathfrak{l}(N)) \subset R(\mathfrak{j})$. By Theorem 6, therefore, $N \subset R(\mathfrak{j})$. Applying the operator $\mathfrak{r}$ to both sides of this inequality, we obtain $\mathfrak{r}(N) \supset \mathfrak{r}(R(\mathfrak{j})) \supset \mathfrak{j}$, using Lemmas 3(a) and 2(a). This, together with $\mathfrak{j} \supset \mathfrak{r}(N)$, yields $\mathfrak{j} = \mathfrak{r}(N)$.

We now prove (2). The totality of automorphisms β for which $\beta\cdot\mathfrak{r}(N) = 0$ forms an ideal $\mathfrak{i}$ which contains $\mathfrak{l}(N)$, so that $\mathfrak{i}\cdot\mathfrak{r}(N) = 0$ and $\mathfrak{i} \supset \mathfrak{l}(N)$. By Lemma 5, $\mathfrak{i}\cdot\mathfrak{r}(N) = 0$ implies $L(\mathfrak{i}) \subset R(\mathfrak{r}(N))$, or $L(\mathfrak{i}) \subset N$ by Theorem 6. Applying $\mathfrak{l}$ to both sides of this inequality and using Lemmas 3(a) and 2(b), we obtain $\mathfrak{i} \subset \mathfrak{l}(L(\mathfrak{i})) \subset \mathfrak{l}(N)$. This, together with $\mathfrak{i} \supset \mathfrak{l}(N)$, gives $\mathfrak{i} = \mathfrak{l}(N)$.

The essential parts of this proof, the steps from $L(\mathfrak{l}(N)) \subset R(\mathfrak{j})$ and $L(\mathfrak{i}) \subset R(\mathfrak{r}(N))$ to $N \subset R(\mathfrak{j})$ and $L(\mathfrak{i}) \subset N$, respectively, are consequences of the basic Theorem 6.

Ideals of the type $\mathfrak{r}(N)$ or $\mathfrak{l}(N)$ are therefore largest ideals which annihilate given ideals on the right or left respectively. Such ideals will be called normal ideals, and we are lead to the following

DEFINITION. An ideal $\mathfrak{r}$ is a *right normal* ideal if there exists an ideal $\mathfrak{a}$ such that $\mathfrak{a}\cdot\mathfrak{r} = 0$ and such that $\mathfrak{a}\cdot\mathfrak{j} = 0$ implies $\mathfrak{j} \subset \mathfrak{r}$. An ideal $\mathfrak{l}$ is a *left normal* ideal if there exists an ideal $\mathfrak{b}$ such that $\mathfrak{l}\cdot\mathfrak{b} = 0$ and such that $\mathfrak{i}\cdot\mathfrak{b} = 0$ implies $\mathfrak{i} \subset \mathfrak{l}$.

The right normal ideal $\mathfrak{r}$ is exactly the ideal consisting of all the automorphisms α for which $\mathfrak{a}\cdot\alpha = 0$, and the left normal ideal $\mathfrak{l}$ of all the automorphisms β for which $\beta\cdot\mathfrak{b} = 0$. Consider now the totality of automorphisms γ which annihilate $\mathfrak{r}$ on the left, $\gamma\cdot\mathfrak{r} = 0$. They form a left normal ideal $\mathfrak{l}$ which contains $\mathfrak{a}$, so that $\mathfrak{l}\cdot\mathfrak{r} = 0$ and $\mathfrak{l} \supset \mathfrak{a}$. The totality of automorphisms α which annihilate $\mathfrak{l}$ on the right, $\mathfrak{l}\cdot\alpha = 0$, is exactly $\mathfrak{r}$ since $\mathfrak{l}\cdot\alpha = 0$ implies $\mathfrak{a}\cdot\alpha = 0$. Hence $\mathfrak{l}\cdot\mathfrak{r} = 0$ and each is the largest ideal which annihilates the other. Likewise, starting with any left normal ideal, there is the corresponding right normal

ideal. We thus have a pairing between right and left normal ideals given by $\mathfrak{l}\cdot\mathfrak{r} = 0$, and each is the largest ideal which annihilates the other.

By Theorem 7, the ideals $\mathfrak{r}(N)$ and $\mathfrak{l}(N)$ are right and left normal ideals respectively, and they are paired as in the preceding paragraph. We shall now prove the basic result that every right or left normal ideal is of the type $\mathfrak{r}(N)$ or $\mathfrak{l}(N)$ respectively.

THEOREM 8. (1) *An ideal* $\mathfrak{r}$ *is of the type* $\mathfrak{r}(N)$ *for some* N *if and only if* $\mathfrak{r}$ *is a right normal ideal.*

(2) *An ideal* $\mathfrak{l}$ *is of the type* $\mathfrak{l}(N)$ *for some* N *if and only if* $\mathfrak{l}$ *is a left normal ideal.*

Proof. We first prove (1). By Theorem 7 it suffices to show that any right normal ideal $\mathfrak{r}$ can be written as $\mathfrak{r}(N)$ for some N. Suppose then that $\mathfrak{a}\cdot\mathfrak{r} = 0$ and that $\mathfrak{a}\cdot\mathfrak{j} = 0$ implies $\mathfrak{j} \subset \mathfrak{r}$. By Lemma 5, $\mathfrak{a}\cdot\mathfrak{r} = 0$ implies that $L(\mathfrak{a}) \subset R(\mathfrak{r})$. Applying the operator $\mathfrak{r}$ to both sides of this inequality and using Lemmas 3(a) and 2(a), we get $\mathfrak{r}(L(\mathfrak{a})) \supset \mathfrak{r}(R(\mathfrak{r})) \supset \mathfrak{r}$. On the other hand, using Lemmas 5 and 2(a), we have $\mathfrak{a}\cdot\mathfrak{r}(L(\mathfrak{a})) = 0$ since $L(\mathfrak{a}) \subset R(\mathfrak{r}(L(\mathfrak{a})))$. Hence $\mathfrak{r}(L(\mathfrak{a})) \subset \mathfrak{r}$ and we must have $\mathfrak{r} = \mathfrak{r}(L(\mathfrak{a}))$. Thus, $\mathfrak{r} = \mathfrak{r}(N)$ for $N = L(\mathfrak{a})$.

In order to prove (2) it suffices to show that any left normal ideal $\mathfrak{l}$ can be written as $\mathfrak{l}(N)$ for some N. Suppose that $\mathfrak{l}\cdot\mathfrak{b} = 0$ and that $\mathfrak{i}\cdot\mathfrak{b} = 0$ implies $\mathfrak{i} \subset \mathfrak{l}$. By Lemma 5, we have $L(\mathfrak{l}) \subset R(\mathfrak{b})$. Applying the operator $\mathfrak{l}$ to both sides of this inequality, we obtain $\mathfrak{l} \subset \mathfrak{l}(L(\mathfrak{l})) \subset \mathfrak{l}(R(\mathfrak{b}))$. On the other hand, $\mathfrak{l}(R(\mathfrak{b}))\cdot\mathfrak{b} = 0$ since $L(\mathfrak{l}(R(\mathfrak{b}))) \subset R(\mathfrak{b})$. Hence $\mathfrak{l} = \mathfrak{l}(R(\mathfrak{b}))$ and we have $\mathfrak{l} = \mathfrak{l}(N)$ for $N = R(\mathfrak{b})$. The theorem is now completely proved.

THEOREM 9. (1) $\mathfrak{r}(R(\mathfrak{r})) = \mathfrak{r}$ *if* $\mathfrak{r}$ *is a right normal ideal;*

(2) $\mathfrak{l}(L(\mathfrak{l})) = \mathfrak{l}$ *if* $\mathfrak{l}$ *is a left normal ideal.*

Proof. By Theorem 8, $\mathfrak{r} = \mathfrak{r}(N)$ for some N. Therefore, $\mathfrak{r}(R(\mathfrak{r})) = \mathfrak{r}(R(\mathfrak{r}(N))) = \mathfrak{r}(N) = \mathfrak{r}$ by Lemma 4(a), and this proves (1).

By Theorem 8, $\mathfrak{l} = \mathfrak{l}(N)$ for some N. Therefore, $\mathfrak{l}(L(\mathfrak{l})) = \mathfrak{l}(L(\mathfrak{l}(N))) = \mathfrak{l}(N) = \mathfrak{l}$ by Lemma 4(b), and (2) is proved.

The major results in this section are contained in Theorems 6, 8 and 9. They show that the operations $R(\mathfrak{r})$ and $\mathfrak{r}(N)$ establish the same one-to-one correspondence between all normal subgroups of A and all right normal ideals of $\mathfrak{o}$ and that the operations $L(\mathfrak{l})$ and $\mathfrak{l}(N)$ establish the same one-to-one correspondence between all normal subgroups of A and all left normal ideals of $\mathfrak{o}$. Furthermore, $\mathfrak{l}(N)\cdot\mathfrak{r}(N) = 0$ and each is the largest which annihilates the other. Finally, it should be remarked that the results of this section all depend on Theorem 6, and that it is only in Theorem 6 that the structure of a normal subgroup N of A is used.

7. The normal ideals. We shall here develop further the theory of §6, and find properties of normal ideals. The symbols $\mathfrak{r}$, $\mathfrak{l}$ shall hereafter stand for right and left normal ideals respectively.

Let $\mathfrak{r}_\nu$ be any set of right normal ideals, $\mathfrak{l}_\nu$ the corresponding left normal ideals, and form the ideals $[\mathfrak{r}_\nu]$, $[\mathfrak{l}_\nu]$, $\{\mathfrak{r}_\nu\}$, $\{\mathfrak{l}_\nu\}$. It is easily seen that $[\mathfrak{r}_\nu]$ (and similarly $[\mathfrak{l}_\nu]$) is a right (left) normal ideal. For, on the one hand $\{\mathfrak{l}_\nu\}\cdot[\mathfrak{r}_\nu] = 0$, and on the other hand any automorphism which annihilates $\{\mathfrak{l}_\nu\}$, $\{\mathfrak{l}_\nu\}\cdot\alpha = 0$, must belong to each $\mathfrak{r}_\nu$ and so to $[\mathfrak{r}_\nu]$. The ideals $\{\mathfrak{r}_\nu\}$ and $\{\mathfrak{l}_\nu\}$, however, are not normal ideals in general. But there exists a smallest right normal ideal which contains $\{\mathfrak{r}_\nu\}$, namely, the intersection of all the right normal ideals containing $\{\mathfrak{r}_\nu\}$. Denote this right normal ideal by $\sum_\nu \mathfrak{r}_\nu$, or by $\mathfrak{r}_1 + \mathfrak{r}_2 + \cdots$. Similarly denote the smallest left normal ideal containing $\{\mathfrak{l}_\nu\}$ by $\sum_\nu \mathfrak{l}_\nu$, or $\mathfrak{l}_1 + \mathfrak{l}_2 + \cdots$. In keeping with this terminology, $[\mathfrak{r}_\nu]$ will be designated by $\prod_\nu \mathfrak{r}_\nu$ or by $\mathfrak{r}_1 \cdot \mathfrak{r}_2 \cdot \cdots$; similarly for $[\mathfrak{l}_\nu]$.

THEOREM 10. (1) *Under the correspondence* $N \leftrightarrow \mathfrak{r}$,

$$\{\ \} \leftrightarrow \prod \quad \textit{and} \quad [\ ] \leftrightarrow \sum.$$

(2) *Under the correspondence* $N \leftrightarrow \mathfrak{l}$,

$$\{\ \} \leftrightarrow \sum \quad \textit{and} \quad [\ ] \leftrightarrow \prod.$$

Proof. (1) states that

$$\mathfrak{r}(\{N_\nu\}) = \prod \mathfrak{r}(N_\nu), \qquad \mathfrak{r}([N_\nu]) = \sum \mathfrak{r}(N_\nu)$$

and

$$R(\prod \mathfrak{r}_\nu) = \{R(\mathfrak{r}_\nu)\}, \qquad R(\sum \mathfrak{r}_\nu) = [R(\mathfrak{r}_\nu)].$$

By Lemma 3, we have the first of these lines with the equality sign replaced by $\subset$ and $\supset$ respectively, and the second line by $\supset$ and $\subset$. Applying the operator R to the first line, one obtains $\{N_\nu\} \supset R(\prod \mathfrak{r}(N_\nu))$ and $[N_\nu] \subset R(\sum \mathfrak{r}(N_\nu))$. These relations, together with the second line, yield the equality sign,

$$\{N_\nu\} = R(\prod \mathfrak{r}(N_\nu)), \qquad [N_\nu] = R(\sum \mathfrak{r}(N_\nu)).$$

This is exactly the second line of our desired result, where $\mathfrak{r}_\nu = \mathfrak{r}(N_\nu)$. Applying $\mathfrak{r}$ to these equations yields the first of the desired lines.

The proof of (2) is similar to that of (1).

Theorem 10 permits carrying over to normal ideals the results obtained for normal subgroups in Part I. Corollary 3 of Part I yields the important distributive law for normal ideals:

THEOREM 11.[13] (1) $\mathfrak{r}_1\cdot\mathfrak{r}_2 + \mathfrak{r}_3 = (\mathfrak{r}_1 + \mathfrak{r}_3)\cdot(\mathfrak{r}_2 + \mathfrak{r}_3)$, $(\mathfrak{r}_1 + \mathfrak{r}_2)\cdot\mathfrak{r}_3 = \mathfrak{r}_1\cdot\mathfrak{r}_3 + \mathfrak{r}_2\cdot\mathfrak{r}_3$.

(2) *Similar to* (1) *for left normal ideals.*

Proof. Apply the operator $\mathfrak{r}$ to the equations in Corollary 3, and use Theorem 10. (1) is obtained in which $\mathfrak{r}_i = \mathfrak{r}(N_i)$ $(i = 1, 2, 3)$.

[13] The extended distributive laws with any number of terms follow from footnote 10.

Theorems 4, 5 of Part I likewise yield important results. Define the normal ideals ${}^h\mathfrak{r}_k$, ${}_n\mathfrak{r}^k$, ${}^h\mathfrak{l}_k$, ${}_n\mathfrak{l}^k$ by

$$ {}^h\mathfrak{r}_k = \mathfrak{r}(p^h A_k), \qquad {}_n\mathfrak{r}^k = \mathfrak{r}({}_nA^k), $$
$$ {}^h\mathfrak{l}_k = \mathfrak{l}(p^h A_k), \qquad {}_n\mathfrak{l}^k = \mathfrak{l}({}_nA^k). $$

Applying the operators $\mathfrak{r}$ and $\mathfrak{l}$ to Corollary 2 shows that ${}_n\mathfrak{r}^k$, ${}^h\mathfrak{l}_k$ are $\sum$-irreducible and ${}^h\mathfrak{r}_k$, ${}_n\mathfrak{l}^k$ are $\prod$-irreducible. Irreducibility is understood in the following sense.

Definition. A right normal ideal $\mathfrak{r}$ is called $\sum$-irreducible if $\mathfrak{r} = \mathfrak{r}_\nu$ for some ν whenever $\mathfrak{r} = \sum_\nu \mathfrak{r}_\nu$. It is called $\prod$-irreducible if $\mathfrak{r} = \mathfrak{r}_\nu$ for some ν whenever $\mathfrak{r} = \prod_\nu \mathfrak{r}_\nu$. Similar definitions are given for irreducible left normal ideals.

Applying the operator $\mathfrak{r}$ (and similarly $\mathfrak{l}$) to Theorems 4, 5 yields the following: any right normal ideal $\mathfrak{r}$ can be represented uniquely in the forms $\mathfrak{r} = \prod_i {}^{h_i}\mathfrak{r}_{n_i+h_i}$ and $\mathfrak{r} = \sum_i {}_{n_i}\mathfrak{r}^{n_i+h_i}$ (the n_i, h_i appearing in both these expressions are not the same). This ideal theory may be summarized in

Theorem 12. (1) *The right normal ideals* ${}_n\mathfrak{r}^k = \mathfrak{r}({}_nA^k)$ *and* ${}^h\mathfrak{r}_k = \mathfrak{r}(p^hA_k)$ *are respectively* $\sum$- *and* $\prod$-*irreducible. Every right normal ideal can be represented uniquely in the form*

$$ \mathfrak{r} = \sum_i {}_{n_i}\mathfrak{r}^{n_i+h_i} $$

where

(a) $0 \leqq n_1 < \cdots < n_i < \cdots, 0 < h_1 < \cdots < h_i < \cdots \leqq \infty$; *and*
(b) $n_i + h_i$ *are relevant orders of* A *for each* i.

Every $\mathfrak{r}$ *can also be represented uniquely in the form*

$$ \mathfrak{r} = \prod_i {}^{h_i}\mathfrak{r}_{n_i+h_i} $$

where

(a) $0 \leqq h_1 < \cdots < h_i < \cdots, 0 < n_1 < \cdots < n_i < \cdots \leqq \infty$; *and*
(b) $n_i + h_i$ *are relevant orders of* A *for each* i.
(2) *Similar to* (1) *for left normal ideals.*

We shall determine when $\sum \mathfrak{r}_\nu$, $\sum \mathfrak{l}_\nu$ coincide with $\{\mathfrak{r}_\nu\}$, $\{\mathfrak{l}_\nu\}$ respectively.

Theorem 13. $\sum_{\nu=1}^{t} \mathfrak{r}_\nu = \{\mathfrak{r}_1, \mathfrak{r}_2, \cdots, \mathfrak{r}_t\}$.

Proof. It suffices to consider the case $t = 2$. Accordingly, let $\mathfrak{r}_1 = \mathfrak{r}(N_1)$, $\mathfrak{r}_2 = \mathfrak{r}(N_2)$ and $\mathfrak{r}_1 + \mathfrak{r}_2 = \mathfrak{r}([N_1, N_2])$ by Theorem 10. Theorem 13 requires proving that any automorphism α for which $[N_1, N_2]\alpha = 0$ can be written in the form $\alpha = \beta + \gamma$, where $N_1\beta = 0$, $N_2\gamma = 0$.

Let $h_1(k)$, $h_2(k)$, $h_3(k)$ denote the invariants of N_1, N_2, $[N_1, N_2]$ respectively, where $h_3(k) = \max h_1(k), h_2(k)$. Let ν' be all those relevant k's for which $h_3(\nu') =$

$h_2(\nu') \geqq h_1(\nu')$; let ν'' be the remaining relevant k's, for which $h_3(\nu'') = h_1(\nu'') > h_2(\nu'')$. Let $b_{g\nu}$ be a complete set of independent primitive elements as in §1. Define β for the basic elements $b_{g\nu}$ by

$$b_{g\nu''}\beta = b_{g\nu''}\alpha, \qquad b_{g\nu'}\beta = 0.$$

This definition of β must be extended to the whole group A.

Suppose that there are only a finite number of ν'', say all $\leqq K$. For any $m \geqq K$, set $A = \sum_{\nu<m} \{b_{g\nu}\} + B^m$ and extend β linearly over A by taking $B^m\beta = 0$. This defines β uniquely, no matter which $m \geqq K$ is taken. If there are only a finite number of ν', interchange the rôles of N_1 and N_2. Consider finally the case when ν' and ν'' occur infinitely often. In particular, this requires that both $h_3(k)$ and $n_3(k)$ tend to ∞ as $k \to \infty$. Let a be any element of A, and choose any number m such that $n_3(m) \geqq n(a)$. Set

$$A = \sum_{\nu<m} \{b_{g\nu}\} + B^m, \qquad a = \sum_{\nu<m} t_{g\nu} b_{g\nu} + c,$$

and define

$$a\beta = \sum_{\nu<m} t_{g\nu} b_{g\nu}\beta = \sum_{\nu''<m} t_{g\nu''} b_{g\nu''}\alpha.$$

The element c has an order $n(c) \leqq n(a) \leqq n_3(m)$; since c belongs to B^m, $h(c) + n(c) \geqq m$ and we have $h(c) \geqq m - n_3(m) = h_3(m)$. Hence $c\alpha = 0$ because $p^{h_3(m)}A_m\alpha = 0$. This shows that the elements $\sum_{\nu''<m} t_{g\nu''}b_{g\nu''}\alpha$ for all m such that $n_3(m) \geqq n(a)$ are identical, and the definition of $a\beta$ is independent of the choice of m. Let a, a' be two elements of A, and select any m such that $n_3(m)$ is greater than both $n(a)$, $n(a')$. Then

$$a' = \sum_{\nu<m} t'_{g\nu}b_{g\nu} + c' \quad \text{and} \quad a + a' = \sum_{\nu<m} (t_{g\nu} + t'_{g\nu})b_{g\nu} + (c + c').$$

Since $n(a + a') \leqq n_3(m)$, it follows immediately that $(a + a')\beta = a\beta + a'\beta$. The automorphism β is completely defined in all cases.

It will now be shown that $N_1\beta = 0$, $N_2(\alpha - \beta) = 0$.

(1) Let $a \subset p^{h_1(\mu')}A_{\mu'}$, i.e., $n(a) \leqq n_1(\mu') = n$ and $h(a) \geqq h_1(\mu') = h$. Choose any sufficiently large m as above. We have

$$a = \sum_{\nu\leqq h+n} t_{g\nu}p^h b_{g\nu} + \sum_{\nu>h+n}^{m-1} t_{g\nu}p^{\nu-n} b_{g\nu} + c,$$

and

$$a\beta = \sum_{\nu''\leqq h+n} t_{g\nu''}p^h b_{g\nu''}\alpha + \sum_{\nu''>h+n}^{m-1} t_{g\nu''}p^{\nu''-n} b_{g\nu''}\alpha.$$

For $\nu'' < \mu'$, $h = h_1(\mu') \geqq h_1(\nu'') = h_3(\nu'')$ so that $p^h b_{g\nu''}\alpha = 0$; for $\nu'' > \mu'$, $n = n_1(\mu') \leqq n_1(\nu'') = n_3(\nu'')$ so that $p^{\nu''-n}b_{g\nu''}\alpha = 0$. Hence $a\beta = 0$, or $p^{h_1(\mu')}A_{\mu'}\beta = 0$. Because $h_3(\mu') = h_2(\mu') \geqq h_1(\mu')$, we also have $p^{h_2(\mu')}A_{\mu'}(\alpha - \beta) = - p^{h_2(\mu')}A_{\mu'}\beta = 0$.

(2) Let $a \subset p^{h_2(\mu'')}A_{\mu''}$, i.e., $n(a) \leqq n_2(\mu'') = n$ and $h(a) \geqq h_2(\mu'') = h$. In addition to the above expressions for a and $a\beta$, we have

$$a\alpha = \sum_{\nu \leqq n+h} t_{\varrho\nu} p^h b_{\varrho\nu}\alpha + \sum_{\nu > n+h}^{m-1} t_{\varrho\nu} p^{\nu-n} b_{\varrho\nu}\alpha$$

since we have already shown that $c\alpha = 0$. For $\nu' < \mu''$, $h = h_2(\mu'') \geqq h_2(\nu') = h_3(\nu')$ so that $p^h b_{\varrho\nu'}\alpha = 0$; for $\nu' > \mu''$, $n = n_2(\mu'') \leqq n_2(\nu') = n_3(\nu')$ so that $p^{\nu'-n}b_{\varrho\nu'}\alpha = 0$. All that remains in the expression for $a\alpha$ are the terms involving ν''; thus $a\alpha = a\beta$, or $p^{h_2(\mu'')}A_{\mu''}(\alpha - \beta) = 0$. Because $h_3(\mu'') = h_1(\mu'') \geqq h_2(\mu'')$ we also have $p^{h_1(\mu'')}A_{\mu''}\beta = p^{h_1(\mu'')}A_{\mu''}(\beta - \alpha) = 0$.

From (1) and (2) above, it follows that $N_1\beta = 0$ and $N_2(\alpha - \beta) = 0$. Setting $\alpha - \beta = \gamma$, we have $\alpha = \beta + \gamma$ where $N_1\beta = 0$, $N_2\gamma = 0$. The proof of the theorem is now complete.

Theorem 14. *If N_1, N_2 are such that $h_{N_1}(k) \leqq h_{N_2}(k)$ for all sufficiently large k, then $\mathfrak{l}(N_1) + \mathfrak{l}(N_2) = \{\mathfrak{l}(N_1), \mathfrak{l}(N_2)\}$.*

Proof. By Theorem 10, it is required to prove that any automorphism α for which $A\alpha \subset \{N_1, N_2\}$ can be written in the form $\alpha = \beta + \gamma$, where $A\beta \subset N_1$, $A\gamma \subset N_2$. Let $h_1(k)$, $h_2(k)$, $h_3(k)$ be the invariants of N_1, N_2, $\{N_1, N_2\}$ respectively. Let ν' be all those relevant k's for which $h_3(\nu') = h_1(\nu') \leqq h_2(\nu')$, and ν'' the remaining relevant k's, for which $h_3(\nu'') = h_2(\nu'') < h_1(\nu'')$. The hypothesis of the theorem states that all ν'' are $\leqq K$.

Set

$$A = \sum_{\nu < K} \{b_{\varrho\nu}\} + B^K,$$

and

$$b_{\varrho'\mu}\alpha = \sum_{\nu < K} t_{\varrho\nu} p^{h_3(\nu)} b_{\varrho\nu} + c,$$

where c belongs to the group generated by $p^{h_3(k)}A_k$ for all $k \geqq K$. Since $h_3(k) = h_1(k)$ for all $k \geqq K$, $c \subset N_1$. Define

$$b_{\varrho'\mu}\beta = \sum_{\nu'' < K} t_{\varrho\nu''} p^{h_3(\nu'')} b_{\varrho\nu''}.$$

We have $b_{\varrho'\mu}\beta \subset N_2$, and

$$b_{\varrho'\mu}(\alpha - \beta) = \sum_{\nu' < K} t_{\varrho\nu'} p^{h_3(\nu')} b_{\varrho\nu'} + c \subset N_1.$$

Also note that $p^K b_{\varrho'\mu}\beta = 0$.

Let a be any element of A of order n, and select any $m \geqq n + K$. We have $a = \sum_{\nu \leqq n} s_{\varrho\nu} b_{\varrho\nu} + \sum_{\nu > n}^{m-1} s_{\varrho\nu} p^{\nu-n} b_{\varrho\nu} + d$, where d belongs to B^m; define

$$a\beta = \sum_{\nu \leqq n} s_{\varrho\nu} b_{\varrho\nu}\beta + \sum_{\nu > n}^{m-1} s_{\varrho\nu} p^{\nu-n} b_{\varrho\nu}\beta.$$

This definition is independent of the m chosen, since $p^{\nu-n}b_{g\nu}\beta = 0$ if $\nu - n \geqq K$. It follows that $(a + a')\beta = a\beta + a'\beta$ as in the proof of Theorem 13.

It is clear that $A\beta \subset N_2$. Furthermore,

$$a(\alpha - \beta) = \sum_{\nu \leqq n} s_{g\nu} b_{g\nu}(\alpha - \beta) + \sum_{\nu > n}^{m-1} s_{g\nu} p^{\nu - n} b_{g\nu}(\alpha - \beta) + d\alpha.$$

The element d belongs to B^m, so that $h(d) + n(d) \geqq m \geqq n + K$; since $n(d) \leqq n$, we have $h(d) \geqq K$. Hence $d\alpha \subset p^K\{N_1, N_2\}$, which is generated by $p^{h_2(k)+K}A_k$ for all $k \geqq K$, and finally $d\alpha \subset N_1$. It follows from the expression for $a(\alpha - \beta)$ that $A(\alpha - \beta) \subset N_1$. The proof of Theorem 14 is complete.

Of importance in connection with (2) of Theorem 12 is the following consequence of Theorem 14.

COROLLARY 6. $\sum_{i=1}^{t} {}^{h_i}\mathfrak{l}_{k_i} = \{{}^{h_1}\mathfrak{l}_{k_1}, \cdots, {}^{h_t}\mathfrak{l}_{k_t}\}$.

The question whether the relation $\mathfrak{l}(N_1) + \mathfrak{l}(N_2) = \{\mathfrak{l}(N_1), \mathfrak{l}(N_2)\}$ holds without any restriction on N_1, N_2 remains unanswered.[14]

Let us note that both $\mathfrak{r}(p^nA)$ and $\mathfrak{l}(A_n)$ consist of all automorphisms α for which $p^n\alpha = 0$, and that these are the only ideals which are both right and left normal ideals.

8. **Concluding remarks.** A similar theory for the group $\mathfrak{G}$ of proper automorphisms of A can be developed. Consider the characteristic subgroups C of A, i.e., those subgroups which remain invariant under all the proper automorphisms α of A.[15] Define the corresponding subgroups of $\mathfrak{G}$:

$\mathfrak{J}(C)$ = invariant subgroup of all proper automorphisms α such that $c\alpha = c$ for each element c of C;

$\mathfrak{K}(C)$ = invariant subgroup of all proper automorphisms β such that $a\beta - a \subset C$ for each a of A.

The inverse operations, defined for every invariant subgroup $\mathfrak{S}$ of $\mathfrak{G}$ are:

$J(\mathfrak{S})$ = totality of elements a such that $a\alpha = a$ for each α in $\mathfrak{S}$;

$K(\mathfrak{S})$ = group generated by elements b of the form $b = a\alpha - a$ for some a of A and α of $\mathfrak{S}$.

It can be shown that $J(\mathfrak{J}(C)) = C$ and $K(\mathfrak{K}(C)) = C$ except for certain types of groups A.[16] But the question of determining which invariant subgroups of $\mathfrak{G}$ are of the type $\mathfrak{J}(C)$ or $\mathfrak{K}(C)$ cannot be answered within the group $\mathfrak{G}$.

The theory for the group $\mathfrak{G}$ can be subsumed under the theory for the ring $\mathfrak{o}$. For, it is immediately seen that $\mathfrak{J}(C)$ and $\mathfrak{K}(C)$ consist of those proper automorphisms of the form $1 + \mathfrak{r}(C)$ and $1 + \mathfrak{l}(C)$ respectively, and that $J(\mathfrak{S}) = \mathfrak{r}(\mathfrak{S}')$, $K(\mathfrak{S}) = \mathfrak{l}(\mathfrak{S}')$, where $\mathfrak{S}'$ is the set of automorphisms of the form $\mathfrak{S} - 1$.

[14] This relation is easily proved for certain types of groups A, e.g., if A is a direct sum of cyclic groups.

[15] See footnote 7.

[16] See Baer, loc. cit. (footnote 3), supplement.

Let $\alpha \subset \mathfrak{J}(C)$, $\beta \subset \mathfrak{K}(C)$; then $\alpha - 1$ and $\beta - 1$ belong to $\mathfrak{r}(C)$ and $\mathfrak{l}(C)$ respectively, and $(\beta - 1)(\alpha - 1) = 0$ or $1 + \beta\alpha = \beta + \alpha$. Thus, an invariant subgroup $\mathfrak{K}$ of $\mathfrak{G}$ is of the type $\mathfrak{K}(C)$ if and only if there is another invariant subgroup $\mathfrak{J}$ of $\mathfrak{G}$ such that $\mathfrak{K}$ consists exactly of all the proper automorphisms β for which $1 + \beta\alpha = \beta + \alpha$ for each α in $\mathfrak{J}$. Similarly for $\mathfrak{J}(C)$. The determination of the invariant subgroups of $\mathfrak{G}$ of the type $\mathfrak{J}(C)$ or $\mathfrak{K}(C)$ therefore requires the operation $+$.

Finally, our theory has immediate application to the study of matrix rings. For example, consider all two-rowed square matrices of the form

$$\begin{pmatrix} r & p^4 s \\ t & u \end{pmatrix},$$

where r, s, t, u are integers taken modulo p^3, p^3, p^3, p^7 respectively. This is isomorphic to the ring of automorphisms of the Abelian group

$$A = \{a\} + \{b\}, \qquad \text{where} \qquad p^3 a = 0,\, p^7 b = 0.$$

The normal subgroups, 20 in number,[17] are determined according to Part I. The right and left normal ideals are then immediately obtained. The ideal structure of these ideals is given by Theorem 12.

NEW YORK UNIVERSITY.

[17] The number of normal subgroups of A has the power of the continuum if the orders of the elements in A have no upper bound, and is $(k_1 + 1) \cdot \prod_{i=2}^{t} (k_i - k_{i-1} + 1)$ if all the relevant orders of A are $k_1, k_2, \cdots, k_t$ in the order of increasing magnitude.

SYLOW THEOREMS FOR INFINITE GROUPS

By Reinhold Baer

There exists a class of results in the theory of finite groups whose proofs are obtained by the method of counting; i.e., one shows that, unless the theorem in question holds true, there is an impossible abundance—or scarcity for that matter—of subsystems with some property. The results obtained this way may be called "Sylow theorems" after their most important representative. Typical examples are the following facts: The existence of central-elements, different from 1, in finite p-groups $\neq 1$; the conjugacy of any two greatest p-subgroups of a finite group; P. Hall's discussion of Sylow-systems of finite soluble groups and so on.

Though trying to break away from the classical limitation to the investigation of finite groups, one may still employ the methods used in the proofs of theorems of the Sylow-type to obtain results concerning groups which are restricted in no way as to size. Certain conditions concerning the subsystems investigated turn out to be needed for the applicability of these methods. They are, however, in general not necessary for the validity of these extensions of theorems of the Sylow-type. But in imposing the condition that finite subsets are contained in finite normal subgroups, a class of groups has been characterized for which a fairly complete theory may be evolved.[1]

Chapter I. Sylow subgroups

1. The only theorem concerning finite groups which we are going to use in this chapter is the *theorem of Cauchy* stating that *a finite group contains an element of order the prime number p if, and only if, the order of the group is divisible by p.* All the other theorems on finite groups, excepting elementary ones, which will be used will be proved as special cases of theorems on groups which may be finite or infinite.

We enumerate some of the notations we are going to use. If S is a subset and x an element of the group G, then $S^x = x^{-1}Sx$. Two subsets U and V of G are termed W-conjugate, where W is also a subset of G, if there exists an element w in W so that $U = V^w$. This relation is symmetric, reflexive and transitive, whenever W is a subgroup of G. U^W signifies the set of all the U^w for w in W.

If S is a subgroup of the group G, then $(S < G)$ is the normalizer of S in G

Received December 13, 1939; presented to the American Mathematical Society, April 12, 1940.

[1] In this connection the following papers ought to be mentioned: A. P. Dietzmann, *Über p-Gruppen*, Comptes Rendus de l'Académie des Sciences de l'URSS, vol. 15(1937), pp. 71–76; A. P. Dietzmann, A. Kurosch, A. I. Uzkow, *Sylow-sche Untergruppen von unendlichen Gruppen*, Recueil Mathématique, vol. 3(1938), pp. 179–184.

which consists of all the elements g in G satisfying $S = S^g$, and $G:S$ is the number of cosets (right or left) of G mod S; i.e., $G:S$ is the index of S in G.

2. If p is a prime number, then the group G is said to be a *p-group* if every element in G is of order a power of p.

THEOREM 2.1. *If G is a p-group, if S is a subgroup of G, and if $G:S$ is finite, then $G:S$ is a power of p.*[2]

Proof. (1) Assume first that S is a normal subgroup of G. Then G/S is a finite group all of whose elements are of order a power of p, since G is a p-group; and it follows from the theorem of Cauchy that the order of G/S, i.e., $G:S$, is a power of p.

(2) Assume secondly that $S = (S < G)$. If w is an element in G which transforms every S^g (for g in G) into itself, then we have $g^{-1}Sg = w^{-1}g^{-1}Sgw$ or $S = (gwg^{-1})^{-1}Sgwg^{-1}$ so that gwg^{-1} belongs to $(S < G) = S$. Hence w belongs to every S^g for g in G. If D is the cross-cut of all the S^g, then D is a normal subgroup of G and G/D is essentially the same as the group of permutations which the elements in G induce in the set of the S^g. Since the number of elements in the set S^G is exactly $G:(S < G) = G:S$, G/D is a finite group. Thus it follows from what has been proved in (1) that $G:D$ is a power of p. Since S/D is a subgroup of G/D, and since $G:S = (G/D):(S/D)$, it follows now that $G:S$ is a power of p. (As a matter of fact, $G = S$, as we shall see later.)

(3) As our theorem certainly holds true in case $G:S = 1$, let us assume that it holds true for all the subgroups T of G so that $G:T < G:S$. We distinguish two cases.

Case 1: $S < (S < G)$. We have in this case $G:(S < G) < G:S$, and it follows from the induction-hypothesis that $G:(S < G)$ is a power of p. Since S is a normal subgroup of $(S < G)$, since $(S < G)$ is a p-group—as a subgroup of G—, it follows from what has been proved in (1) that $(S < G):S$ is a power of p; and hence $G:S$ is a power of p, since $G:S = [G:(S < G)][(S < G):S]$.

Case 2: $S = (S < G)$. In this case it follows from what has been proved in (2) that $G:S$ is a power of p.

THEOREM 2.2. *If G is a p-group, if S is a subgroup of G which is not a normal subgroup of G, if the set S^G of conjugate subgroups to S is finite, then there exists a subgroup T of G which is different from S, conjugate to S and satisfies* $(S < G) = (T < G)$.

Proof. The number of elements in S^G is certainly $G:(S < G)$. Hence it follows from Theorem 2.1 that the number of elements in S^G is a power of p and is actually divisible by p, since it is $\neq 1$.

If T is any subgroup in the set S^G, then the set $T^{(S<G)}$ is a subset of S^G and therefore finite and two sets $T^{(S<G)}$ and $T'^{(S<G)}$ are either equal or disjoint.

[2] See Dietzmann, Kurosch, Uzkow, op. cit., p. 182.

If T^* is the cross-cut of $(S < G)$ and $(T < G)$, then the number of elements in $T^{(S<G)}$ is exactly $(S < G):T^*$; and this number is a power of p by Theorem 2.1. Since $(S < G):S^* = 1$, it follows that there exists at least one $T \neq S$ in S^G so that $(S < G):T^* = 1$. This implies $(S < G) \leqq (T < G)$; and since S and T are conjugate subgroups of G, it follows that $(S < G) = (T < G)$.

The following statement is a simple consequence of this theorem:

COROLLARY 2.3. *If G is a p-group, if S is a subgroup of G so that $S < G$ and so that $G:(S < G)$ is finite,*[3] *then $S < (S < G)$.*

COROLLARY 2.4.[4] *If G is a p-group, if S is a subgroup of G so that $S < G$ and so that $G:(S < G)$ is finite, then there exists a normal subgroup N of G so that $S \leqq N < G$.*

Proof. This statement is certainly true if $G:(S < G) = 1$, since then S itself is a normal subgroup of G. Thus we may assume that it holds true for every subgroup T of G so that $T < G$ and $G:(T < G) < G:(S < G)$. If $1 < G:(S < G)$, then it follows from the finiteness of $G:(S < G)$ and from Corollary 2.3 that $(S < G) < ((S < G) < G)$ so that $G:((S < G) < G) < G:(S < G)$. Hence it follows from the induction-hypothesis that there exists a normal subgroup N of G so that $(S < G) \leqq N < G$; and this completes the proof.[5]

3. A subgroup P of the group G is said to be a *p-component of G* if P is a p-group which is not a proper subgroup of any p-group contained in G. Since the identity is a p-group, one may verify as usual that there exist p-components of every group.

THEOREM 3.1.[6] *If there exists one p-component P of G so that P^G is a finite set, then any two p-components of G are conjugate subgroups of G and the number of p-components of G is* $\equiv 1 \pmod p$.

Proof. Denote by Q some p-component of G; and denote by W the set of subgroups in the finite set P^G which are different from Q. If R is some p-component in W, then R and Q are different. Since Q and R are both p-components of G, it is therefore impossible that one is a subgroup of the other. Thus there exists an element u in Q that is not contained in R. If $R = u^{-1}Ru$, then the subgroup, generated by u and R, would be a p-group, since u is an element of

[3] I.e., the set S^G of subgroups of G which are conjugate to S in G is finite. Cf. Theorem 3.6 below.

[4] Cf. Theorem 3.7 below.

[5] The proof that a finite p-group, different from 1, possesses a central, different from 1, may be used to prove the following statement. If G is a p-group, and if F is a finite and normal subgroup of G which is different from 1, then F contains an element, different from 1, which is contained in the central of G. Cf. the paper of Dietzmann (see footnote 1); and R. Baer, *Nilpotent groups and their generalizations*, Trans. Amer. Math. Soc., vol. 47(1940), pp. 393–434.

[6] The first proof of this theorem has been given by Dietzmann, Kurosch and Uzkow in their paper cited in footnote 1. Our proof is rather different from theirs and is just an adaptation of the proof reproduced in H. Zassenhaus, *Lehrbuch der Gruppentheorie*.

order a power of p. But this is impossible, since R is a p-component. Thus it follows that the set R^Q does not consist of one element only. Since $R \neq Q$, it follows that Q is not an element in R^Q so that $R^Q \leqq W$. If R^* is the cross-cut of Q and $(R < G)$, then it follows now that $Q:R^*$ is the number of elements in R^Q. Since Q is a p-group, it follows from Theorem 2.1 that the number of elements in R^Q is a power of p which is different from 1; and thus it follows finally that the number of subgroups in W is $\equiv 0 \pmod p$. If we apply this result first on the special case $Q = P$, we see that the number of subgroups in P^G is $\equiv 1 \pmod p$. If there were a p-component V not contained in P^G, applying the above result to $Q = V$, we should find that the number of elements in P^G is $\equiv 0 \pmod p$. This being a contradiction, our theorem is proved.

Remark 3.2. It has been pointed out[7] that the condition of the theorem cannot be omitted without invalidating the theorem. That the condition is, however, not a necessary one may be seen from the following example:[8]

Denote by Z_i for every integer $i \gtreqless 0$ a cyclic group of order 2; and by H the free product of all these groups. There exists clearly one and only one automorphism $\mathbf{g}$ of H so that $Z_{i-1}^{\mathbf{g}} = Z_i$. Let G be the group generated in adjoining to H an element y so that $y^{-1}xy = x^{\mathbf{g}}$ for every x in H.

Clearly H contains all the elements of finite order in G. It is a consequence of Kurosch's theorem[9] on the subgroups of free products that the 2-components of H are just the subgroups Z_i and those subgroups which are conjugate to some Z_i. In G, however, all the Z_i are conjugate so that all the 2-components of G are conjugate, though their number is infinite.

Corollary 3.3. *If P is a p-component of G, and if $G:P$ is finite, then $G:P$ is relatively prime to p.*

Proof. Clearly $G:P = [G:(P < G)]\cdot[(P < G):P]$. That the first factor is relatively prime to p is a consequence of Theorem 3.1. That the second factor is prime to p may be proved as follows. P is a normal subgroup of $(P < G)$; and it is at the same time a p-component of $(P < G)$. If $(P < G):P$ were divisible by p, then it would follow from Cauchy's theorem that $(P < G)/P$ contains an element Pw of order p. Its representative w is clearly of order a power of p. Since w transforms P into itself, it follows now that a p-group is generated in adjoining w to P; and this is impossible.

Corollary 3.4. *If P is a p-component of the finite group G, then the order of P is the highest power of p dividing the order of G.*

Proof. That the order of P is a power of p is a consequence of Cauchy's theorem and the fact that P is a p-group. The other statements are now a consequence of Corollary 3.3.

Remark. This last corollary may be stated in the following form: If G is a

[7] Cf. the paper of Dietzmann, Kurosch and Uzkow, cited in footnote 1.

[8] Another example can be found in the paper of Dietzmann, Kurosch and Uzkow.

[9] A. Kurosch, *Die Untergruppen der freien Produkte von beliebigen Gruppen*, Math. Ann., vol. 109(1934), pp. 647–660.

finite group, then every p-component is a Sylow subgroup and conversely. (Sylow theorem.)

COROLLARY 3.5. *If P is a p-component of G so that the number of subgroups of G which are conjugate to P is finite, and if q is a prime number which does not divide the number of subgroups conjugate to P, i.e., if q and $G:(P < G)$ are relatively prime, then there exists in every class of conjugate q-components of G one which is a subgroup of $(P < G)$.*

Proof. Suppose that Q is some fixed q-component of G. Then the sets P^{gQ} for g in G are disjoint and finite (as subsets of the finite set P^G). A set P^{gQ} contains $Q:(P^g)^*$ elements, where $(P^g)^*$ is the meet of Q and $(P^g < G)$; and this number is by Theorem 2.1 a power of q, i.e., either 1 or $\equiv 0 \pmod q$. If this number were not 1 for any of the P^g, then $G:(P < G)$ would be $\equiv 0 \pmod q$; and this is impossible from our hypothesis. Hence there exists at least one element g in G so that P^{gQ} consists of one element only, i.e., so that $Q \leqq (P^g < G)$. But then $Q^{g^{-1}} \leqq (P < G)$; and this is our statement.

Since every finite positive integer is divisible by but a finite number of prime numbers, the above corollary implies that $(P < G)$ *contains q-components of G for almost every prime number q, if $G:(P < G)$ is finite.*

THEOREM 3.6. *The group G is a direct product of p-groups if, and only if,*

(a) *every element in G is of finite order;*

(b) *there exists for every prime number p a p-component P of G so that the number of conjugates to P in G is finite;*

(c) *every subgroup S of G whose index $G:S$ is finite and different from 1 is a proper subgroup of its normalizer $(S < G)$ in G.*

Proof. If G is a direct product of p-groups, then G is the direct product of uniquely determined p-components; and this implies (a) and (b). If S is a subgroup of such a group G, then S is the direct product of its subgroups S_p, where S_p is the meet of S and G_p, and where G_p is the set of all the elements of order a power of p in G. If $G:S$ is finite, then $G:S$ is the product of the numbers $G_p:S_p$. If $G:S \neq 1$ and is finite, then every $G_p:S_p$ is finite and only a finite number, though at least one, of the numbers $G_p:S_p$ is different from 1. It is a consequence of Corollary 2.3 that $G_p:S_p \neq 1$, though finite, implies that $S_p < (S_p < G_p)$. Since $(S < G)$ is the direct product of the $(S_p < G_p)$, this implies the necessity of condition (c).

Assume now conversely that the conditions (a) to (c) are satisfied. If P is a p-component, then P is the only p-component of $(P < G)$ and we have therefore $(P < G) = ((P < G) < G)$. If P possesses but a finite number of conjugates in G, then $G:(P < G)$ is finite and it follows from (c) that $G = (P < G)$, i.e., P is a normal subgroup of G. Hence P contains every element of order a power of p, since by their adjunction to P a p-group is generated; and P is therefore the only p-component of G. It follows from (b) that there exists for every prime number p one and only one p-component G_p and that G_p contains

all the elements of order a power of p; and it follows from (a) that G is the direct product of the G_p.

THEOREM 3.7. *The group G is a direct product of p-groups if, and only if,*

(a) *every element in G is of finite order;*

(b) *there exists for every prime number p a p-component P of G so that the number of conjugates to P in G is finite;*

(c) *there exists to every proper subgroup S of G whose index $G:S$ is finite a normal subgroup N of G so that $S \leqq N < G$.*

Proof. Denote by G_p the set of all the elements of order a power of p in G. If G is a direct product of p-groups, then every G_p is a subgroup of G and G is the direct product of the G_p. Conditions (a), (b) are satisfied as in Theorem 3.6. If S is a subgroup of G, then S is the direct product of the meets S_p of S and G_p. If there exists one prime number p so that $G_p:S_p$ is finite and different from 1, then it follows from Corollary 2.4 that there exists a normal subgroup N_p of G_p so that $S_p \leqq N_p < G_p$. If N is the direct product of N_p and of the G_q for $q \neq p$, then N is a normal subgroup of G and $S \leqq N < G$; and this proves the necessity of condition (c).

Suppose conversely that the conditions (a) to (c) are satisfied by the group G. It is a consequence of (b) and of Theorem 3.1 that any two p-components of G are conjugate in G. Denote by P some p-component of G. If P were not a normal subgroup of G, then $(P < G) < G$, though $G:(P < G)$ is finite (by (b)). Hence it follows from (c) that there exists a normal subgroup N of G so that $(P < G) \leqq N < G$. Let w be any element in G that is not contained in N. Then $P^w \leqq N^w = N$ so that both P and P^w are p-components of N. Since the number of p-components of G is finite, so is the number of p-components of N; and it follows from Theorem 3.1 that any two p-components of N are conjugate in N. Hence there exists an element v in N so that $P = P^{wv}$, and this implies that wv is an element in $(P < G)$ and therefore in N. As v is in N, this would imply that w is in N; and we have a contradiction. Hence P is a normal subgroup of G; and it follows from (a) that G is the direct product of its p-components.[10]

4. A group is said to be *locally finite*,[11] if every finite subset is contained in a finite and normal subgroup. Examples of locally finite groups, containing a given number of elements, are easily constructed. If a p-group is locally finite, then it can be proved that it is nilpotent. The converse does not hold true, as may be seen from the following example of a *group with Abelian central quotient-group which is not locally finite though all of its elements $\neq 1$ are of order p.*

[10] Further criteria for representability of a group as a direct product of p-groups may be found in the author's paper mentioned in footnote 5.

[11] This concept is different from the concept "local endlich" introduced by Dietzmann, Kurosch, Uzkow (op. cit., footnote 1).

This group G is generated by elements u_i, v_i, w for $i = 1, 2, \cdots$ which are subject to the following relations:

$$1 = u_i^p = v_i^p = w^p;$$

$$u_iu_j = u_ju_i, \quad u_iv_j = v_ju_i, \quad u_iw = wu_i, \quad v_iv_j = v_jv_i;$$

$$v_i^{-1}wv_i = wu_i.$$

It is clear that the central of G contains all the elements u_i so that the central quotient-group of G is Abelian. This implies in particular that all the elements in G are of order 1 or p, if $p \neq 2$; and if $p = 2$, then the orders of the elements are divisors of 4. A normal subgroup of G which contains w contains all the u_i and is therefore infinite.

Theorem 4.1. *A subgroup P of the locally finite group G is a p-component of G if, and only if, the following condition is satisfied.*

(*) *The cross-cut of P and F is a p-component of F for every finite normal subgroup F of G.*

Proof. Suppose first that condition (*) is satisfied by P. If x is any element in P, then there exists a finite normal subgroup X of G which contains x, since G is locally finite. The element x is contained in the cross-cut of P and X; and this cross-cut is a p-group by (*). Hence x is of order a power of p and P is a p-group. If P is a subgroup of the p-group Q, and if w is an element in Q, then let W be a finite and normal subgroup of G which contains w. As Q is a p-group, it follows that the cross-cut of W and Q is a p-group. Since (*) is satisfied by P, and since $P \leqq Q$, it follows that P and Q have the same cross-cut with W. Hence w is an element in P, i.e., $P = Q$ and P is a p-component of G.

Assume now conversely that P is a p-component of the locally finite group G. In order to prove that (*) is satisfied by P, we prove first the following statement.[12]

(4.1.1) *If N is a normal subgroup of the finite group M, and if K is a p-component of M, then the cross-cut of N and K is a p-component of N.*

Denote by L the cross-cut of N and K. Then L is contained in a p-component L^* of N, and L^* is contained in some p-component H of M. Since M is finite, it follows from Corollary 3.4 that H and K are conjugate subgroups of M so that there exists an element g in M satisfying: $H^g = K$. Since N is a normal subgroup of M, we have $N = N^g$. Since H is a p-group and L^* a p-component of N, it follows that L^* is the cross-cut of H and N. Hence $L^{*g} = L$ and this proves that L is a p-component of N.

(4.1.2) *If F is a finite and normal subgroup of the locally finite group G, if the subgroup S of G is a p-group, then there exists a subgroup T of G which is a p-group, contains S, and whose cross-cut with F is a p-component of F.*

[12] This result, of course, is well known. Cf., however, Corollary 4.2 below.

It suffices to prove the existence of a p-component T^* of F which contains the cross-cut S^* of S and F and whose join T with S is a p-group.

Assume first that V is a p-component of F which contains S^* though its join with S is not a p-group. Then there exists a finite subset U of S so that the join of V and U is not a p-group. Since G is locally finite, there exists a finite and normal subgroup F' of G which contains both U and V. Since U is part of the cross-cut of F' and S, it follows in particular that the join of V and the cross-cut of S and F' is not a p-group.

Assume now that every p-component V of F which contains S^* has the property that its join with S is not a p-group. Since F is a finite group, there exists but a finite number of p-components of F. Since G is locally finite, it follows that the join of a finite number of finite, normal subgroups of G is a finite and normal subgroup of G. Hence it follows from what has been proved in the preceding paragraph that there exists a finite and normal subgroup W of G with the following properties: $F \leqq W$; if W^* is the cross-cut of W and S, and if V is a p-component of F which contains S^*, then the join of V and W^* is not a p-group.

There exists, however, a p-component W^{**} of W which contains W^*. It is a consequence of (4.1.1) that the cross-cut V^{**} of W^{**} and F is a p-component of F, since W is finite and F is a normal subgroup of W. Clearly $S^* \leqq V^{**}$. But this is a contradiction. Hence (4.1.2) holds true.

That every p-component of a locally finite group satisfies (*) is a fairly obvious consequence of (4.1.2).

The following statement is a generalization of (4.1.1).

COROLLARY 4.2. *If N is a normal subgroup of the locally finite group G, and if P is a p-component of G, then the cross-cut of P and N is a p-component of N.*

Proof. Denote by P' the cross-cut of P and N, and by F a finite and normal subgroup of N. Since F need not be a normal subgroup of G, we imbed F into a finite, normal subgroup of G whose cross-cut F^* with N is a finite and normal subgroup of G too, since N is a normal subgroup of G. The cross-cut P'' of P and F^* is by Theorem 4.1 a p-component of F^*. Since F is a normal subgroup of F^*, it follows from (4.1.1) that the cross-cut of P'' and F, which is just the meet of P' and F, is a p-component of F. Since every subgroup of a locally finite group is itself locally finite, it follows now from Theorem 4.1 that P' is a p-component of N.

Suppose that N is a normal subgroup of the group G, and that $\mathbf{g}$ is an automorphism of N. Then $\mathbf{g}$ is said to be a *normal* automorphism of N, provided $S = S^{\mathbf{g}}$ for every subgroup S of N which is at the same time a normal subgroup of G (and therefore of N). Every inner automorphism of G induces normal automorphisms in the normal subgroups. But the converse does not hold true.

The normal automorphisms of locally finite groups G are exactly those which induce normal automorphisms in all those subgroups of G which are finite and normal at the same time, since G is the join of its finite and normal subgroups.

Thus the normal automorphisms of the finite, normal subgroups will be of particular importance. It may be remarked that normal automorphisms of the normal subgroup N of G induce normal automorphisms in every subgroup M of N which is normal in G.

We are going to discuss *properties* $\mathbf{P}$ of normal automorphisms of the finite and normal subgroups which are subject to the following condition:

(4.P) *If H and $K \leqq H$ are both finite and normal subgroups of G, and if the normal automorphism $\mathbf{g}$ of H obeys property $\mathbf{P}$, then the automorphism which $\mathbf{g}$ induces in K satisfies $\mathbf{P}$.*

We shall have to discuss in the course of our investigation several instances of such properties $\mathbf{P}$.

THEOREM 4.3. *If G is a locally finite group, if $\mathbf{P}$ is a property of normal automorphisms of normal, finite subgroups of G which obeys the rule* (4.P), *and if there exists corresponding to every finite, normal subgroup N of G a normal automorphism of N which satisfies $\mathbf{P}$, then there exists a normal automorphism of G which induces in every finite, normal subgroup of G an automorphism, satisfying $\mathbf{P}$.*

Proof. It will be convenient to introduce the following notation. If S is a normal subgroup of G, then an automorphism $\mathbf{f}$ of S is termed *admissible*, provided $\mathbf{f}$ satisfies the following conditions.

(i) $\mathbf{f}$ is normal.

(ii) If F is a finite subgroup of S and a normal subgroup of G, then $\mathbf{P}$ is satisfied by the automorphism which $\mathbf{f}$ induces in F.

Using this terminology, we set out to prove the existence of admissible automorphisms of the group G.

Since G is locally finite, there exists an ascending chain of normal subgroups G_ν of G with the following properties.

(1) $G_0 = 1$.

(2) $G_{\nu+1}/G_\nu$ is a finite group.

(3) G_ν is for limit-ordinals ν the set of all the elements contained in groups G_μ for $\mu < \nu$.

(4) $G_\gamma = G$.

We are now going to construct by complete (transfinite) induction admissible automorphisms $\mathbf{g}(\nu)$ of G_ν so that $\mathbf{g}(\rho)$ and $\mathbf{g}(\sigma)$ coincide on G_ρ whenever $\rho < \sigma$. For the purposes of our inductive procedure it is necessary to impose a further condition upon these automorphisms.

(iii) If F is a finite and normal subgroup of G, then there exists an admissible automorphism of F which coincides with $\mathbf{g}(\nu)$ on the cross-cut of F and G_ν.

It is a consequence of the hypothesis of our theorem that $\mathbf{g}(0) = 1$ meets the requirements (i) to (iii). Thus we may assume that $\mathbf{g}(\mu)$ has been defined for every $\mu < \nu$ in such a way as to meet all our requirements.

Case 1. $\nu = \tau + 1$. There exists by (2) a finite and normal subgroup N of G so that G_ν is the join of N and G_τ. It is a consequence of the hypothesis of our theorem that there exists at least one admissible automorphism of N;

and as N is a finite group, there exists at most a finite number of admissible automorphisms of N. Now we prove the following statement.

(4.3.1) There exists an admissible automorphism $\mathbf{v}$ of N with the following property.

(+) If H is a finite and normal subgroup of G, and if $N \leqq H$, then there exists an admissible automorphism $\mathbf{h}$ of H which coincides with $\mathbf{v}$ on N and with $\mathbf{g}(\tau)$ on the meet of H and G_τ.

If this statement were not true, there would exist corresponding to every admissible automorphism $\mathbf{f}$ of N a finite and normal subgroup $H = H(\mathbf{f})$ of G for which condition (+) would not be fulfilled though $N \leqq H$. The join W of all the $H(\mathbf{f})$ is a finite and normal subgroup of G, since there exists but a finite number of admissible automorphisms of N, and since G is locally finite. Hence it follows from the induction-hypothesis that we may apply (iii) upon W, G_τ, and $\mathbf{g}(\tau)$. Consequently there exists an admissible automorphism $\mathbf{w}$ of W which coincides with $\mathbf{g}(\tau)$ on the meet of W and G_τ. Since W contains all the $H(\mathbf{f})$, and since every $H(\mathbf{f})$ contains N, it follows that N is a subgroup of W, and that therefore $\mathbf{w}$ induces an admissible automorphism $\mathbf{w}'$ in N. Since $\mathbf{w}$ is an admissible automorphism of W, it induces an admissible automorphism $\mathbf{w}''$ in the subgroup $H(\mathbf{w}')$ of W. This admissible automorphism $\mathbf{w}''$ coincides with $\mathbf{w}'$ on N and with $\mathbf{g}(\tau)$ on the meet of $H(\mathbf{w}')$ and G_τ, since $\mathbf{w}$ coincides with $\mathbf{g}(\tau)$ on the meet of W and G_τ. But this contradicts our choice of $H(\mathbf{w}')$ so that our assumption of the falsity of (4.3.1) has led to a contradiction.

Denote now by $\mathbf{v}$ a fixed admissible automorphism of N which satisfies (+). Then we may prove the following statement.

(4.3.2) Suppose that D and E are any two finite and normal subgroups of G which are contained in G_τ. Denote by D' and E' the joins of N and D and of N and E respectively; and denote by $\mathbf{d}$ and $\mathbf{e}$ admissible automorphisms of D' and E' respectively which coincide with $\mathbf{v}$ on N and with $\mathbf{g}(\tau)$ on the cross-cut of G_τ and D' or E' respectively. Then $\mathbf{d}$ and $\mathbf{e}$ coincide on the meet of D' and E'.

Denote by U the join of D' and E'. Since N is a subgroup of U, and since U is a finite and normal subgroup of the locally finite group G, there exists by (+) an admissible automorphism $\mathbf{u}$ of U which coincides with $\mathbf{v}$ on N and with $\mathbf{g}(\tau)$ on the meet of G_τ and U. This implies that $\mathbf{u}$, $\mathbf{d}$ and $\mathbf{e}$ coincide on N (with $\mathbf{v}$), that $\mathbf{u}$ and $\mathbf{d}$ coincide on D (with $\mathbf{g}(\tau)$), and that $\mathbf{u}$ and $\mathbf{e}$ coincide on E (with $\mathbf{g}(\tau)$). Thus $\mathbf{u}$ and $\mathbf{d}$ coincide on D'; and $\mathbf{u}$ and $\mathbf{e}$ coincide on E'; and this implies our proposition.

If x is any element in G_ν, then there exists a finite and normal subgroup X of G which is contained in G_τ so that x is an element of the join of X and N. There exists by (+) an admissible automorphism of the join of X and N which coincides with $\mathbf{v}$ on N and with $\mathbf{g}(\tau)$ on X. It is a consequence of (4.3.2) that all these admissible automorphisms, independent of the particular choice of X, map the element x upon the same element $x^{\mathbf{g}(\nu)}$. If x and y are any two elements in G_ν, then there exists a finite and normal subgroup Y of G which is part

of G_ν so that both x and y are elements of the join of Y and N. It is a consequence of (4.3.2) that $\mathbf{g}(\nu)$ induces an automorphism of this join-group; and this implies that $\mathbf{g}(\nu)$ is an automorphism. In the same way one sees that $\mathbf{g}(\nu)$ is an admissible automorphism of G_ν.

Suppose now that R is any finite and normal subgroup of G. Then the join K of R and N is a finite and normal subgroup of G. There exists a finite and normal subgroup D of G which is contained in G_τ so that the join of D and N contains the cross-cut of K and G_ν. If H is the join of K and D, then H is a finite and normal subgroup of G and the meet of H and G_ν is generated by D and N. There exists by (+) an admissible automorphism $\mathbf{f}$ of H which induces $\mathbf{v}$ in N and which coincides with $\mathbf{g}(\tau)$ on the cross-cut of H and G_τ. Since $\mathbf{f}$ therefore coincides with $\mathbf{g}(\tau)$ on D, it follows that $\mathbf{f}$ coincides with $\mathbf{g}(\nu)$ on the meet of H and G_ν; and this shows finally that (iii) is satisfied by the admissible automorphism $\mathbf{g}(\nu)$ of G_ν which clearly induces $\mathbf{g}(\tau)$ in G_τ.

Case 2. ν is a limit-ordinal. In this case G_ν is the set of all the elements contained in groups G_μ for $\mu < \nu$. There exists therefore one and only one automorphism $\mathbf{g}(\nu)$ of G_ν which coincides with $\mathbf{g}(\mu)$ on G_μ for every $\mu < \nu$. It is readily seen that this automorphism is admissible and satisfies (iii).

Thus we have succeeded in constructing the automorphisms $\mathbf{g}(\nu)$ for every ν. Hence there exists in particular the admissible automorphism $\mathbf{g}(\gamma)$ of $G_\gamma = G$; and this completes the proof of our theorem.

THEOREM 4.4. *If G is a locally finite group, and if R and S are two p-components of G, then there exists a normal automorphism of G which maps R upon S.*

Proof. Suppose that F is a finite and normal subgroup of G. Then a normal automorphism of F is said to satisfy property $\mathbf{P}$ if it maps the cross-cut of F and R upon the cross-cut of F and S. This property $\mathbf{P}$ satisfies clearly condition (4.P).

Since G is locally finite, it follows from Theorem 4.1 that the cross-cut of F and R as well as the cross-cut of F and S are both p-components of F. Since F is finite, there exists (by Sylow's theorem) an inner automorphism of F which maps the cross-cut of F and R upon the cross-cut of F and S. Thus it follows that the hypothesis of Theorem 4.3 is satisfied by the property $\mathbf{P}$. Hence there exists by Theorem 4.3 a normal automorphism $\mathbf{f}$ of G with the following property.

If F is a finite and normal subgroup of G, then the cross-cut of F and R is mapped by $\mathbf{f}$ upon the cross-cut of F and S.

Since, however, every element in G is contained in some finite and normal subgroup F of G, it follows that R is mapped upon S by $\mathbf{f}$; and this completes the proof.

THEOREM 4.5. (a) *Any two p-components of the countable,*[13] *locally finite group G are conjugate subgroups of G if, and only if, there exists only a finite number of p-components of G.*

[13] Whether or not this countability-hypothesis is necessary for the validity of Theorem 4.5, the author has not been able to decide.

(b) *The following statements are satisfied by every countable and locally finite group G.*

(b′) *If the number of p-components of G is finite, then any two p-components of G are conjugate subgroups of G.*

(b″) *If there exists an infinity of p-components of G, then every complete set of conjugate p-components contains a countable infinity of elements, and there exist $2^{\aleph_0}$ different classes of conjugate p-components of G.*

Proof. (a) is an obvious consequence of (b). In order to prove (b) we prove first the following statement.

(4.5.1) *If G is countable and locally finite, and if there exists an infinity of p-components of G, then there exist $2^{\aleph_0}$ p-components of G.*

Since G is countable and locally finite, there exists an ascending chain of finite and normal subgroups $N(i)$ of G so that every element of G is contained in at least one of the groups $N(i)$ $(i = 0, 1, 2, \cdots)$, $N(0) = 1$.

It is a consequence of (4.1.1) that every p-component of $N(i + 1)$ meets $N(i)$ in a p-component; and every p-component of $N(i)$ is contained in some p-components of $N(i + 1)$. If H and K are two p-components of $N(i)$, then there exists, by Corollary 3.4, an element g in $N(i)$ so that $H^g = K$; and the inner automorphism, induced by g in $N(i + 1)$, maps the set of p-components of $N(i + 1)$ which meet $N(i)$ in H upon the set of those p-components of $N(i + 1)$ which meet $N(i)$ in K. Thus there exists a finite positive integer $f(i)$ with the property: corresponding to every p-component H of $N(i)$ there exist exactly $f(i)$ p-components of $N(i + 1)$ which meet $N(i)$ in H.

If P is any p-component of G, then it follows from Theorem 4.1 that the cross-cut $P(i)$ of P and $N(i)$ is a p-component of $N(i)$. These subgroups $P(i)$ form clearly an ascending chain of subgroups so that every element in P is contained in at least one $P(i)$.

Suppose conversely that $P(i)$ is, for every i, a p-component of $N(i)$, that $P(i) \leqq P(i + 1)$, and that P is the subgroup of all the elements which are contained in at least one $P(i)$. Since every finite and normal subgroup of G is contained in at least one subgroup $N(j)$, it follows from (4.1.1) and from Theorem 4.1 that P is a p-component of G. This shows that the number of p-components of G is exactly the number of ascending chains of subgroups $P(i)$ so that $P(i)$ is a p-component of $N(i)$. This number is, however, easily computed. For, if all the $f(i)$ for $k < i$ are equal to 1, then the number is $f(1) \cdots f(k)$ and is therefore finite. If it happens on the other hand for an infinity of i that $f(i) \neq 1$, then there exist $2^{\aleph_0}$ different such chains, i.e., there exist $2^{\aleph_0}$ different p-components of G. This completes the proof of (4.5.1).

To prove (b) we recall first that G is countable so that all complete classes of conjugate subgroups are countable. If one complete class of conjugate p-components of G is finite, then it is a consequence of Theorem 3.1 that any two p-components of G are conjugate. Now (b) is an obvious consequence of (4.5.1).

Remark. It is comparatively easy to construct countably infinite and locally finite groups whose p-components are not conjugate. Let $B(i)$ be some finite

group, containing more than i p-components. The direct product of the $B(i)$ is countably infinite, locally finite and contains an infinity of p-components. It follows from Theorem 4.5 that not any two of its p-components are conjugate.

Chapter II. Sylow bases

5. The subgroups S_p of G are said to constitute a *Sylow-basis* of G, if they satisfy the following conditions:

(i) S_p is a p-component of G.

(ii) G is generated by the S_p.

(iii) If x is an element in the subgroup generated by $S_{p_1}, \cdots, S_{p_k}$, then its order is an integer, not divisible by prime numbers different from $p_1, \cdots, p_k$.

Thus only groups without elements of infinite order may possess a Sylow-basis. Hall[14] has proved the following fundamental theorems concerning Sylow-bases of finite groups.

(E) *There exists a Sylow-basis of the finite group G if, and only if, G is a soluble group.*

(C) *Any two Sylow-bases of a finite group G are conjugate in G.*

(S) *Any given Sylow-basis of any subgroup S of the finite soluble group G consists of the cross-cuts of S with the members of a suitably chosen Sylow-basis of G.*

The proofs in this chapter are largely effected by reduction to the finite case, whereas we avoided this method throughout the first chapter.

6. This section is devoted to a generalization of the theorem (C) of §5.

THEOREM 6.1. *If there exists one Sylow-basis of the group G which possesses but a finite number of different conjugate Sylow-bases in G, then any two Sylow-bases are conjugate in G.*

Proof. Suppose that the p-components S_p (for every prime number p) constitute a Sylow-basis of G which possesses but a finite number of different conjugates in G. If S is the set of these S_p, g an element in G, then we denote by S^g the set of the S_p^g. As there exists but a finite number of distinct S^g, there exists only a finite number of different S_p^g for every p; and it follows from Theorem 3.1 that *any two p-components of G are conjugate subgroups of G.* This implies in particular that every set S_p^G consists of all the p-components of G.

An element u in G transforms every S^g into itself if, and only if, it transforms every S_p^g into itself, i.e., if, and only if, it transforms all the p-components of G (for every p) into themselves. If U is the set of all the elements u with this property that $u^{-1}S^gu = S^g$ for every g in G, then U is a normal subgroup of G, G/U is essentially the same as a group of permutations of the set of all the S^g so that G/U is a finite group, since the set of the S^g is finite.

Suppose now that T is some p-component of G, and that T^* is the subgroup of G/U which is represented by elements in T. Then T^* is a p-group, and

[14] P. Hall, *On the Sylow systems of a soluble group*, Proc. London Math. Soc., vol. 43(1937), pp. 316–323.

there exists therefore a p-component W^* of G/U which contains T^*. If $T^* \neq W^*$, it would follow from Corollary 2.3 that there exists an element w^* in W^* with the following properties: $T^* = T^{*w^*}$ and w^{*p} is an element in T^*, though w^* is not in T^*. Since all the elements in G are of finite order, there exists an element w of order a power of p so that $w^* = Uw$. If t is an element in T, then it follows from $T^* = T^{*w^*}$ that $w^{-1}tw = ut'$ for u in U and t' in T. As both u and t' transform T into itself, it follows that $w^{-1}tw$ is an element of order a power of p that transforms T into itself. Since T is a p-component of G, this implies that $w^{-1}tw$ is an element in T, i.e., $T = T^w$. Since w is therefore an element of order a power of p, transforming the p-component T into itself, it follows that w belongs to T; and this contradicts our original choice of w. Hence we have proved that *T represents a p-component of G/U.*

Assume now that B is a Sylow-basis of G, consisting of the p-components B_p, that B_p^* is the subgroup of G/U which is represented by the elements in B_p, and that B^* is the set of the B_p^*. If P is some (finite) set of prime numbers, then denote by $B(P)$ the subgroup of G which is generated by the B_p for p in P, and denote by $B^*(P)$ the subgroup of G/U which is generated by the B_p^* for p in P. Then $B^*(P)$ is exactly the subgroup of G/U which is represented by elements in $B(P)$ and it follows that *B^* is a Sylow-basis of G/U.*

Now it follows from (C) of §5 that S^* and B^* are conjugate in G/U, i.e., there exists an element v^* in G/U so that $S_p^{*v^*} = B_p^*$ for every prime number p. If v is an element in G so that $v^* = Uv$, t an element in S_p, then $v^{-1}tv = ub$ for u in U and b in B_p; and it follows as above that $v^{-1}tv$ belongs to B_p, since it is an element of order a power of p that transforms the p-component B_p into itself. Hence $S_p^v = B_p$ for every p, i.e., $S^v = B$; and this completes the proof.

Note that we have proved the following statement.

COROLLARY 6.2. *If B is a Sylow-basis of the group G, if U consists of all the elements in G which transform every B^g for g in G into itself, then U is a normal subgroup of G and every Sylow-basis of G represents a Sylow-basis of G/U.*

The following statement is easily derived from this corollary.

COROLLARY 6.3. *If the group G possesses a Sylow-basis B so that B^G is a finite set, and if S_p is a p-component of G, then there exists only a finite number of prime numbers q so that the normalizer of S_p in G does not contain every q-component of G.*

This statement is an improvement upon Corollary 3.5 and its consequences under the additional hypothesis of the existence of a Sylow-basis possessing but a finite number of conjugates.

7. In this section we are going to investigate the Sylow-bases of locally finite groups.

LEMMA 7.1. *The subgroups S_p of the locally finite group G form a Sylow-basis of G if, and only if, the meets of the S_p and N constitute a Sylow-basis of N for every finite and normal subgroup N of G.*

Proof. If the S_p form a Sylow-basis of the locally finite group G, and if N is a finite and normal subgroup of G, then it is a consequence of Theorem 4.1 that the meet of S_p and N is a p-component of N. Hence it follows from Corollary 3.4 that N is generated by the meets of N and the S_p. If finally P is a (finite) set of prime numbers, then the join $S(P)$ of the S_p for p in P contains only elements whose orders are not divisible by prime numbers not in P; and consequently the same condition is satisfied by the meets of N and S_p. Thus a Sylow-basis of N is formed by the meets of N and S_p.

Assume conversely that the set S_p of subgroups of G satisfies the condition of this lemma. Then it is a consequence of Theorem 4.1 that every S_p is a p-component of G. Since every element in G is contained in some finite and normal subgroup of G, it follows that G is generated by the S_p. If finally P is a (finite) set of prime numbers, $S(P)$ is the subgroup generated by the S_p for p in P, and t is an element in $S(P)$, then there exists a finite set F of elements so that every element in F belongs to some S_p for p in P, and so that t is contained in the subgroup generated by F. Thus the smallest normal subgroup of G, containing F, is a finite and normal subgroup N of G, containing F and t; and now it follows from our hypothesis that the order of t is not divisible by prime numbers, not contained in P. Hence a Sylow-basis of G is formed by the S_p.

COROLLARY 7.2. *If the subgroups S_p of the locally finite group G form a Sylow-basis of G, and if N is some normal subgroup of G, then a Sylow-basis of N is formed by the meets of N and the S_p.*

Proof. We note first that N is locally finite, since the meets of N and finite, normal subgroups of G are finite and normal subgroups of N. Suppose now that H is a finite and normal subgroup of N. Then H is contained in a finite and normal subgroup L of G. It is a consequence of Lemma 7.1 that the meets L_p of L and S_p form a Sylow-basis of L. The cross-cut K of N and L is a normal subgroup of L. Hence it follows that the meets K_p of L_p and K form a Sylow-basis of K. Since finally H is a normal subgroup of K, it follows that the meets H_p of K_p and H form a Sylow-basis of H. But H_p is the meet of H and S_p; and now it follows that the meets of S_p and N form (by Lemma 7.1) a Sylow-basis of N.

THEOREM 7.3. *If G is a locally finite group, then each of the following properties implies the others*:

(a) *There exists a Sylow-basis of G.*

(b) *If N is a finite and normal subgroup of G, then there exists a Sylow-basis of N.*

(c) *Every finite and normal subgroup of G is soluble.*

Proof. Since the equivalence of (b) and (c) is a consequence of (E) in §5, we are concerned only with the equivalence of (a) and (b). That (b) is a consequence of (a) follows immediately from Lemma 7.1. Thus we assume finally that (b) is satisfied by the locally finite group G.

There exists an ascending chain of subgroups G_ν of G so that

(i) $G_0 = 1$;

(ii) every G_ν is a normal subgroup of G;

(iii) $G_{\nu+1}/G_\nu$ is a finite group;

(iv) G_ν is for limit-ordinals ν the set of all the elements contained in groups G_μ for $\mu < \nu$;

(v) $G_\gamma = G$.

Then we are going to construct by complete transfinite induction Sylow-bases $S_{p\nu}$ of G_ν so that the $S_{p\mu}$ are for $\mu < \nu$ just the meets of G_μ with $S_{p\nu}$. Since $S_{p0} = 1$, we may assume that we have already succeeded in constructing the Sylow-basis $S_{p\mu}$ for $\mu < \nu$ which meet our requirements.

Case 1. $\nu = \tau + 1$. In this case there exists a finite and normal subgroup N of G so that G_ν is the join of G_τ and N. It is a consequence of our hypothesis that there exists a Sylow-basis of N; and since N is finite, there exists a finite number of distinct Sylow-bases of N. Let T_p be some Sylow-basis of N; and denote by T_p^* the join of T_p and $S_{p\tau}$. Every finite and normal subgroup of G_ν is contained in the join of N and some finite and normal subgroup of G which is contained in G_τ. Hence it follows from Lemma 7.1 and Corollary 7.2 that the T_p^* form a Sylow-basis of G_ν, if a Sylow-basis of N' is formed by the meets of N' and the T_p^*, whenever N' is the join of N and some finite, normal subgroup of G which is contained in G_τ.

Suppose now that none of the possible systems T_p^* forms a Sylow-basis of G_ν. Then there exists corresponding to each Sylow-basis T_p of N a finite and normal subgroup W of G which is contained in G_τ so that the meets of T_p^* with the join of W and N do not form a Sylow-basis of this join. The subgroup V of G which is generated by all these subgroups W is finite, since there exists but a finite number of Sylow-bases of N; V is furthermore normal in G and part of G_τ. It is a consequence of Corollary 7.2 that the meets V_p of V and $S_{p\tau}$ form a Sylow-basis of V. If U is the join of N and V, then it follows from (S) of §5 that there exists a Sylow-basis U_p of U so that the V_p are just the meets of V and U_p. These U_p meet N in a Sylow-basis N_p; and the U_p are just the meets of U and N_p^* (= join of N_p and $S_{p\tau}$). But this contradicts our choice of V; and thus it follows that the Sylow-basis $S_{p\tau}$ of G_τ may be extended to a Sylow-basis of G_ν.

Case 2. ν is a limit-ordinal. Then the joins $S_{p\nu}$ of the $S_{p\mu}$ for $\mu < \nu$ meet all the requirements.

Thus the $S_{p\nu}$ have been defined for every ν. Since the $S_{p\gamma}$ form a Sylow-basis of $G_\gamma = G$, this completes the proof.

This criterion may be put into a different form. A group G has been termed *metacyclic*[15] if there exists an ascending chain of subgroups G_ν, satisfying the following conditions.

(i) $G_0 = 1$.

[15] Cf. R. Baer, op. cit., footnote 5.

(ii) G_ν is a normal subgroup of $G_{\nu+1}$ and $G_{\nu+1}/G_\nu$ is a cyclic group.

(iii) G_ν is for limit-ordinals ν the set of all the elements contained in groups G_μ for $\mu < \nu$.

(iv) $G_\gamma = G$.

Subgroups and quotient-groups of metacyclic groups are metacyclic; and finite metacyclic groups are the same as finite soluble groups. Thus it follows that locally finite groups are metacyclic if, and only if, each of its finite and normal subgroups is soluble. The conditions (a) to (c) of Theorem 7.3 are therefore equivalent to the following property of locally finite groups.

(d) *G is metacyclic.*

THEOREM 7.4. *If G is a locally finite group, and if the subgroups S_{pi} of G form (for $i = 1, 2$) a Sylow-basis of G, then there exists a normal automorphism* $\mathbf{f}$ *of G so that $S_{p2} = S_{p1}^{\mathbf{f}}$ for every prime number p.*

Proof. If F is a finite and normal subgroup of G, then the normal automorphism $\mathbf{g}$ of F is said to have property $\mathbf{P}$ if it maps the meet of S_{p1} and F upon the meet of S_{p2} and F for every prime number p. This property $\mathbf{P}$ satisfies condition (4.P). It is furthermore a consequence of Lemma 7.1 that the meets of F and S_{pi} form for $i = 1$ or 2 a Sylow-basis of F; and it follows therefore from (C) of §5 that there exist normal (even inner) automorphisms of F which satisfy $\mathbf{P}$. Hence it follows from Theorem 4.3 that there exists a normal automorphism $\mathbf{f}$ of G which satisfies $S_{p2} = S_{p1}^{\mathbf{f}}$ for every prime number p, since every finite set of elements in the locally finite group G is contained in some finite and normal subgroup F of G.

THEOREM 7.5. (a) *Any two Sylow-bases of the countable, locally finite group G are conjugate in G if, and only if, there exists only a finite number of different Sylow-bases of G.*

(b) *The following properties are satisfied by countable and locally finite groups G:*

(b′) *If the number of Sylow-bases of G is finite, then any two Sylow-bases of G are conjugate in G.*

(b″) *If there exists an infinity of Sylow-bases of G, then every complete set of conjugate Sylow-bases contains a countable infinity of different Sylow-bases; and there exist $2^{\aleph_0}$ different classes of conjugate Sylow-bases of G.*

The proof of this theorem is constructed on essentially the same lines as the proof of Theorem 4.5.

THE UNIVERSITY OF ILLINOIS.

LAGUERRE POLYNOMIALS AND THE LAPLACE TRANSFORM

By J. Shohat

Introduction. The object of this paper is to study the Laplace transform

$$F(t) = \int_0^\infty e^{-tu}\phi(u)\,du, \qquad R(t) > t_0 > 0, \tag{1}$$

where

$$\int_0^\infty e^{-u}\,|\,\phi(u)\,|^2\,du$$

exists, by means of the orthonormal Laguerre polynomials

$$\phi_n(u) = \phi_n(u;0,\infty;e^{-u}) = \frac{1}{n!}\left(u^n - \frac{n^2}{1!}u^{n-1} + \frac{n^2(n-1)^2}{2!}u^{n-2} + \cdots\right),$$

$$\int_0^\infty e^{-u}\phi_n(u)\phi_m(u)\,du = \delta_{m,n} \qquad (m, n = 0, 1, 2, \cdots). \tag{2}$$

The connection between the Laplace integral and Laguerre polynomials has been exhibited by Widder [15][1] who made use of it in order to obtain an inversion formula for the general Laplace transform

$$\int_0^\infty e^{-tu}\,d\alpha(u) \qquad (\alpha(u) \text{ bounded, non-decreasing}). \tag{3}$$

Here we are concerned mainly with the uniqueness of the representation (1) and with the nature of $F(t)$. Our discussion is based upon the now classical property of Laguerre polynomials expressed in Parseval's formula

$$\int_0^\infty e^{-u}f_1(u)f_2(u)\,du = \sum_{n=0}^\infty A_nB_n,$$

where

$$A_n = \int_0^\infty e^{-u}f_1(u)\phi_n(u)\,du, \qquad B_n = \int_0^\infty e^{-u}f_2(u)\phi_n(u)\,du.$$

(4), where the series converges absolutely, holds for any two functions $f_{1,2}(u)$ such that $\int_0^\infty e^{-u}\,|\,f_{1,2}(u)\,|^2\,du$ exists[2] (which implies the existence of

Received January 1, 1940; presented to the American Mathematical Society, December 31, 1936.

[1] Numbers in brackets refer to the bibliography at the end.

[2] $\phi(u)$ in (1) and also $f(u), f_1(u), \cdots$ are in general complex-valued functions of the real variable u; e.g., $f_1(u) = \psi_1(u) + i\psi_2(u)$, where the ψ's are real.

$\int_0^\infty e^{-u} f_{1,2}(u) u^n \, du \ (n = 0, 1, 2, \cdots)\big)$. (4) yields, in particular,

$$\int_0^\infty e^{-u} |f(u)|^2 du = \sum_{n=0}^{\infty} \left| \int_0^\infty e^{-u} f(u) \phi_n(u) \, du \right|^2. \tag{5}$$

The use of Parseval's formula enables us to establish *the equivalence of the functional representation* (1) *to the class of functions belonging to* (H_2), *that is, representable in the region* D: $|z| < 1$ *by power series* $\sum_{n=0}^{\infty} c_n z^n$, *with* $\sum_{n=0}^{\infty} |c_n|^2 < \infty$. The same considerations, combined with certain results from the theory of functions of a complex variable, enable us to treat the problem of *uniqueness* of the representation (1). Moreover, they yield many results from the theory of approximation—generalizations of Weierstrass' Theorem obtained in entirely different ways by Szász [12] and Müntz [7].

Our procedure may be characterized thus: from uniqueness of the Laplace transform to approximation, while most writers on this subject use the reverse procedure—from approximation to the said uniqueness [6, 16].[3]

1. Writing in (1)

$$\int_0^\infty e^{-tu} \phi(u) \, du = \int_0^\infty e^{-(t-\frac{1}{2})u - \frac{1}{2}u} \phi(u) \, du,$$

we see that the existence of $\int_0^\infty e^{-u} |\phi(u)|^2 du$ implies $t_0 = \frac{1}{2}$, i.e.,

$$R(t) > \tfrac{1}{2}. \tag{6}$$

We wish to apply (4) to the right member in (1). To this end, rewrite (1) as

$$F(t) = \int_0^\infty e^{-u} \cdot e^{-(t-1)u} \phi(u) \, du. \tag{7}$$

By virtue of (2), we get at once

$$\int_0^\infty e^{-u} \cdot e^{-(t-1)u} \phi(u) \, du = \frac{1}{t}\left(\frac{1}{t} - 1\right)^n \qquad (n = 0, 1, 2, \cdots), \tag{8}$$

$$e^{-(t-1)u} \sim \sum_{n=0}^{\infty} \frac{1}{t}\left(\frac{1}{t} - 1\right)^n \phi_n(u).$$

Parseval's formula now yields the *fundamental relation*

$$\begin{aligned} F(t) &= \int_0^\infty e^{-tu} \phi(u) \, du = \frac{1}{t} \sum_{n=0}^{\infty} A_n \left(\frac{1}{t} - 1\right)^n, \qquad R(t) > \tfrac{1}{2}, \\ A_n &= \int_0^\infty e^{-u} \phi_n(u) \phi(u) \, du \qquad \left(\int_0^\infty e^{-u} |\phi(u)|^2 du \text{ exists}\right). \end{aligned} \tag{9}$$

[3] Cf. also a recent paper by W. Feller [2], also the work of Picone [9], Hille [3], Hille and Tamarkin [4], Paley and Wiener [8] and Tricomi [14], which have points in common with the present paper. The author is indebted to E. Hille for calling his attention to [3, 4, 8, 9, 14].

The series in (9) converges absolutely (the convergence of the series in (8) does not concern us here).

We could write down (8) and (9) directly, making use of the generating function for Laguerre polynomials

$$\frac{1}{1+z}e^{zu/(1+z)} \sim \sum_{n=0}^{\infty} \phi_n(u)z^n,$$

where we set

$$z = \frac{1}{t} - 1; \qquad |z| < 1 \tag{10}$$

(by (6)). (9) now becomes

$$F(t) \equiv F\left(\frac{1}{1+z}\right) = (1+z)\sum_{n=0}^{\infty} A_n z^n, \qquad |z| < 1, \tag{11}$$

where, by (5),

$$\sum_{n=0}^{\infty} |A_n|^2 = \int_0^{\infty} e^{-u}|\phi(u)|^2 du < \infty. \tag{12}$$

It follows that *the functions*

$$w(z) = \sum_{n=0}^{\infty} A_n z^n, \qquad w_1(z) = (1+z)w(z) = \sum_{n=0}^{\infty}(A_{n-1}+A_n)z^n, \tag{13}$$

$$|z| < 1 \qquad\qquad (A_{-1} = 0)$$

both belong to (H_2).

Conversely, *given the power series* $\sum_{n=0}^{\infty} a_n z^n$, *with* $\sum_{n=0}^{\infty} |a_n|^2 < \infty$, *so that it represents a function* $A(z) \epsilon (H_2)$, $(1+z)A(z)$ *can be represented as a Laplace integral* (1), *namely,*

$$(1+z)A(z) = \int_0^{\infty} e^{-tu}f(u)\,du, \tag{14}$$

$$\int_0^{\infty} e^{-u}|f(u)|^2 du \textit{ exists}; \qquad t = \frac{1}{1+z}, \qquad R(t) > \tfrac{1}{2}.$$

By hypothesis,

$$a_n = c_n + id_n, \qquad \sum_{n=0}^{\infty} c_n^2 < \infty, \qquad \sum_{n=0}^{\infty} d_n^2 < \infty.$$

By the Riesz-Fischer Theorem, there exist (real) functions $\psi_1(u)$, $\psi_2(u)$ such that

$$\psi_1(u) \sim \sum_{n=0}^{\infty} c_n\phi_n(u), \qquad \psi_2(u) \sim \sum_{n=0}^{\infty} d_n\phi_n(u),$$

$$\int_0^{\infty} e^{-u}\psi_1^2(u)\,du = \sum_{n=0}^{\infty} c_n^2, \qquad \int_0^{\infty} e^{-u}\psi_2^2(u)\,du = \sum_{n=0}^{\infty} d_n^2.$$

The desired representation (14) is now obtained if we let

$$f(u) = \psi_1(u) + i\psi_2(u),$$

and observe that, by virtue of (4) and (10), where $|z| < 1$ implies $R(t) > \frac{1}{2}$ and vice versa,

$$(1+z)A(z) = \frac{1}{t}\sum_{n=0}^{\infty} a_n\left(\frac{1}{t}-1\right)^n = \int_0^{\infty} e^{-u}\cdot e^{-(t-1)u}f(u)\,du = \int_0^{\infty} e^{-tu}f(u)\,du.$$

We have thus established the equivalence of the two representations as described in the introduction.

We conclude that in (1) $F(1/(1+z))$ possesses all properties of functions $\epsilon\ (H_2)$ ([13], p. 244).

Remark. The relation (11) may be considered as an inversion formula for (1) in the following restricted sense. $F(t)$ being given, (11) enables us to find the coefficients A_n in the expansion $\phi(u) \sim \sum_{n=0}^{\infty} A_n\phi_n(u)$, and the set $\{A_n\}$ determines $\phi(u)$ uniquely (disregarding nul-functions); $\phi(u)$ may now be determined almost everywhere in $(0, \infty)$ as the limit in the mean of the sequence

$$s_n(u) = \sum_{i=0}^{n} A_i\phi_i(u) \qquad (n = 0, 1, 2, \cdots).$$

2. The question as to the uniqueness of the representation (1) evidently reduces to the following one. *If in* (1) $F(t) = \int_0^{\infty} e^{-tu}\phi(u)\,du$ *vanishes for a certain infinite sequence of distinct values of* t, *can we conclude that* $\phi(u)$ *vanishes almost everywhere in* $(0, \infty)$? This, in turn, implies the vanishing of all A_n in (11), hence the vanishing in D of the functions $w(z)$, $w_1(z)$ as given in (13). To abbreviate, we introduce the following

DEFINITION I. $\phi(u)$, for which $\int_0^{\infty} e^{-u}\,|\phi(u)|^2\,du$ exists, is said to satisfy condition $(0\text{-}z_k)$ if the equations

$$F(t_k) \equiv F\left(\frac{1}{1+z_k}\right) = (1+z_k)w(z_k) = \sum_{n=0}^{\infty} A_n z_k^n, \qquad |z_k| < 1$$

$$(z_k \neq z_l,\ k \neq l;\ l, k = 1, 2, \cdots)$$

(see (11)) imply $\phi(u) = 0$ almost everywhere in $(0, \infty)$, in particular, at all points of continuity (that is, uniqueness of the representation (1)).

THEOREM I. (i) *Condition* $(0\text{-}z_k)$ *is satisfied if the set* $\{z_k = t_k^{-1} - 1\}$ $(z_k \neq 1,\ k \neq l,\ z_k \neq z_l)$ *has a limit point inside the circle* $|z| = 1$.

(ii) *If* $\lim_{k\to\infty} |z_k| = 1$, *condition* $(0\text{-}z_k)$ *is still satisfied, provided the function* $w_1(z) = \sum_{n=0}^{\infty} A'_n z^n = (1+z)w(z)$ *is bounded in* D: $|z| < 1$ *and* $\prod_{k=1}^{\infty} |z_k| = 0$.

(iii) *If* $\lim_{k=\infty} |z_k| = 1$, *the condition* ($0$-$z_k$) *does not always hold, if* $\prod_{k=1}^{\infty} |z_k| > 0$.

First consider (i). Here $w(z)$, hence $(1 + z)w(z)$, vanishes identically in D, and this implies, by virtue of (11) and (5), that $\int_0^\infty e^{-u} |\phi(u)|^2 du = 0$.

Next we prove (ii). Here again, $w_1(z)$, being bounded in D, vanishes therein identically, under the conditions stated ([1], p. 122), and the same conclusion as in (i) follows.[4]

Finally, consider (iii). Construct a function $\Phi(z)$ regular-analytic and bounded in D, hence of the class (H_2) ([5], pp. 25–26), which vanishes at all points z_k and yet does not vanish identically in D ([1], loc. cit.).[5] The same properties belong to the function.

$$(15) \qquad (1 + z)\Phi(z) = \sum_{n=0}^{\infty} c_n z^n \equiv F(t), \qquad t = \frac{1}{1 + z}, \qquad |z| < 1,$$

and by the preceding discussion,

$$F(t) = \int_0^\infty e^{-tu}\phi(u)\, du, \qquad \phi(u) \sim \sum_{n=0}^{\infty} c_n\phi_n(u),$$

$$\int_0^\infty e^{-u}|\phi(u)|^2 du = \sum_{n=0}^{\infty} |c_n|^2 > 0.$$

It follows that the vanishing almost everywhere in $(0, \infty)$ of $\phi(u)$ is impossible, although $F(t_k) = 0$ $(k = 1, 2, \cdots)$.

Remark. The function $w_1(z) \equiv F(1/(1 + z))$ in (1) is certainly bounded in D, if $|\phi(u)| \leqq M$ in $(0, \infty)$, where M is a fixed number. In general, the condition that a function $\sum_{n=0}^{\infty} B_n z^n$ be bounded in D can be expressed either in the form of non-negativeness of certain determinants involving the B_n ([1], p. 145) or in the following form ([5], p. 17):

$$\frac{1}{n + 1}\left|\sum_{k=0}^{n}\sum_{i=0}^{k} B_i z^i\right| < \text{fixed number } B, \quad |z| = 1 \qquad (n = 0, 1, \cdots).$$

We now let

$$t_k = x_k + iy_k, \qquad |z| = 1 - v_k \qquad (k = 1, 2, \cdots).$$

The relation $1 + z_k = t_k^{-1}$ gives

$$(16) \qquad \frac{2x_k - 1}{x_k^2 + y_k^2} < v_k < 2\cdot\frac{2x_k - 1}{x_k^2 + y_k^2} \qquad (x_k > \tfrac{1}{2};\ 0 < v_k < 1).$$

[4] This conclusion holds even for less restricted $w_1(z)$. (Cf., e.g., P. Dienes, *The Taylor Series*, p. 237.) However, boundedness suffices for the subsequent applications.

[5] In Bieberbach's example $\Phi(z) = \prod_{k=1}^{\infty} (z - z_k)/(z - w_k)$, where $z_k = r_k e^{i\phi_k}$, $0 < \delta < r_k < 1$, $\lim_{k=\infty} r_k = 1$, we take $w_k = (2 - r_k)e^{i\phi_k}$. Then all conditions required are satisfied. Moreover, $|(z - z_k)/(z - w_k)| < 1$, $k \geqq 1$, so that $|\Phi(z)| \leqq 1$, for $|z| < 1$.

Hence, the two series

$$\sum_{k=1}^{\infty} v_k \equiv \sum_{k=1}^{\infty} (1 - |z_k|), \qquad \sum_{k=1}^{\infty} \frac{2x_k - 1}{x_k^2 + y_k^2}$$

converge or diverge together. Moreover, the convergence of $\prod_{k=1}^{\infty} |z_k|$ is equivalent to that of $\sum_{k=1}^{\infty} v_k$. This leads to

THEOREM II. *The conclusions of Theorem* I (ii), (iii) *hold, under the conditions stated therein, if the quantities* $t_k = x_k + iy_k$ $(x_k > \frac{1}{2};\ k = 1, 2, \cdots)$ *are such that* $\sum_{k=1}^{\infty} (2x_k - 1)/(x_k^2 + y_k^2)$ (ii) *diverges,* (iii) *converges respectively. In particular, the conclusion in* (iii) *holds if* $\sum_{k=1}^{\infty} x_k^{-1}$ *converges.*

The substitution $e^{-u} = y$ enables us to extend the foregoing considerations to integrals of the form

$$f(\mu) = \int_0^1 y^{\mu} \psi(y)\, dy; \qquad \int_0^1 |\psi(y)|^2\, dy \text{ exists.} \tag{17}$$

Here again we say that $\psi(y)$, *for which* $\int_0^1 |\psi(y)|^2$ *exists, satisfies "condition* $(0\text{-}\mu_k)$" *if the vanishing of* $f(\mu)$ *in* (17) *for a given infinite set* $\{\mu_k\}$ *implies the vanishing of* $\psi(y)$ *almost everywhere in* (0, 1).

THEOREM III. $\psi(y)$ *in* (17) *satisfies condition* $(0\text{-}\mu_k)$, *where*

$$\mu_k = \xi_k + i\eta_k, \qquad \xi_k > -\tfrac{1}{2} \qquad (k = 1, 2, \cdots),$$

provided that the sequence of points $\{\xi_k = (\mu_k + 1)^{-1} - 1\}$ *has a limit point inside the circle* $|\xi| = 1$. *The condition* $(0\text{-}\mu_k)$ *is not always satisfied if* $\lim_{k=\infty} |\xi_k| = 1$ *and* $\sum_{k=1}^{\infty} (2\xi_k + 1)/[(\xi_k + 1)^2 + \eta_k^2]$ *converges, in particular, if* $\sum_{k=1}^{\infty} (\xi_k + 1)^{-1}$ *converges.*

3. We now turn to the theory of approximation.

DEFINITION II. An infinite sequence of numbers

$$\mu_k = \xi_k + i\eta_k\ (k = 0, 1, \cdots), \qquad \mu_0 = 0; \qquad \xi_k > 0,\ k \geqq 1, \tag{18-C}$$

or

$$\mu_k = \xi_k + i\eta_k\ (k = 0, 1, 2, \cdots); \qquad \xi_k > -\tfrac{1}{2},\ k \geqq 0, \tag{18-L}$$

is said to be a *base* for $C(0, 1)$ or $L_2(0, 1)$, respectively, if to any preassigned $\epsilon > 0$ there corresponds a linear aggregate $c_n(x) = \sum_{k=0}^{n} c_k x^{\mu_k} + c_k' x^{\bar{\mu}_k}$, with con-

stant coefficients, such that

$$\text{(19-C)} \qquad | f(x) - c_n(x) | < \epsilon \qquad (0 \leqq x \leqq 1),$$

$$\text{(19-L)} \qquad \int_0^1 | f(x) - c_n(x) |^2 \, dx < \epsilon$$

(x^{μ_k}-principal value; $x^0 \equiv 1$ for $0 \leqq x \leqq 1$), for any $f(x) \in C(0, 1)$ or $\in L_2(0, 1)$ respectively.[6]

When discussing various ways of approximating functions as in (19-C, L), the following lemma proves useful.

LEMMA. (i) *If the sequence* (18-C) *is a base for* $C(0, 1)$, *it is also a base for* $L_2(0, 1)$ (*also for* $L_p(0, 1)$, $p \geqq 1$).

(ii) *If the sequence* (18-L) *is a base for* $L_2(0, 1)$, *the sequence*

$$\nu_{-1} = 0, \qquad \nu_k = \mu_k + 1, \qquad \bar{\nu}_k \qquad (k = 0, 1, 2, \cdots)$$

is a base for $C(0, 1)$.

(i) is almost self-evident. Its proof, along customary lines, is as follows. If $f(x) \in L_2(0, 1)$, a continuous function $\phi(x)$ exists such that

$$\int_0^1 | f(x) - \phi(x) |^2 \, dx < \tfrac{1}{2}\epsilon.$$

To this $\phi(x)$ corresponds, by hypothesis, an aggregate $c_n(x)$, as specified above, such that

$$| \phi(x) - c_n(x) | < (\tfrac{1}{2}\epsilon)^{\frac{1}{2}} \qquad (0 \leqq x \leqq 1).$$

Hence,

$$\int_0^1 | f(x) - c_n(x) |^2 \, dx \leqq 2 \left\{ \int_0^1 | f(x) - \phi(x) |^2 \, dx + \int_0^1 | \phi(x) - c_n(x) |^2 \, dx \right\} < \epsilon.$$

We first prove statement (ii) for a function $f(x)$ having in (0, 1) a continuous derivative. Then

$$\text{(20)} \qquad \begin{aligned} & f(x) - f(0) - \sum_{k=0}^{\infty} c_k x^{\mu_k + 1} + c_k' x^{\bar{\mu}_k + 1} \\ & = \int_0^x [f'(u) - \sum_{k=0}^{n} c_k(\mu_k + 1) u^{\mu_k} + c_k'(\bar{\mu}_k + 1) u^{\bar{\mu}_k}] \, du \qquad (0 \leqq x \leqq 1), \end{aligned}$$

$$\text{(21)} \qquad \begin{aligned} & | f(x) - f(0) x^{\nu_{-1}} - \sum_{k=0}^{n} c_k x^{\nu_k} + c_k' x^{\bar{\nu}_k} |^2 \\ & \leqq \int_0^1 | f'(u) - \sum_{k=0}^{n} c_k \nu_k u^{\mu_k} + c_k' \bar{\nu}_k u^{\bar{\mu}_k} |^2 \, du, \end{aligned}$$

[6] This means $f(x) = \psi_1(x) + i\psi_2(x)$, $\psi_{1,2}(x) \in C(0, 1)$ or $\in L_2(0, 1)$ respectively. We introduce here explicitly the conjugate $x^{\bar{\mu}_k}$ for the following reason. $| f(x) - \sum_{k=0}^{n} A_k x^{\mu_k} | < \epsilon$ in (0, 1) implies $| \bar{f}(x) - \sum_{k=0}^{n} \bar{A}_k x^{\bar{\mu}_k} | < \epsilon$, also $| f(x) - \sum_{k=0}^{n} \frac{1}{2}(A_k x^{\mu_k} + \bar{A}_k x^{\bar{\mu}_k}) | < \epsilon$ in (0, 1), if $f(x)$ is real-valued.

where, by hypothesis, the right-hand term can be made $< \epsilon^2$, by a proper choice of n and the constants c_k, c'_k. We proceed to extend our result to $f(x)$ assumed only to be continuous in $(0, 1)$. First, introduce the function $f_1(x)$ defined as follows:

$$f_1(x) = f(x) \qquad (0 \leqq x \leqq 1),$$

$$f_1(x) = f(1), \qquad 1 \leqq x \leqq 1 + b \qquad (b > 0, \text{ arbitrarily fixed}).$$

$f_1(x)$ being continuous in $(0, 1 + b)$, choose $h > 0$ and $< b$ so that

$$| f_1(x + h) - f_1(x) | < \epsilon \qquad (0 \leqq x \leqq x + h \leqq 1 + b).$$

With h thus fixed, we now introduce "Stekloff's function" [11]

$$F(x) = \frac{1}{h}\int_x^{x+h} f_1(y)\,dy \qquad (0 \leqq x < x + h \leqq 1 + b). \tag{22}$$

Then

$$| F(x) - f_1(x) | \leqq \frac{1}{h}\int_x^{x+h} | f_1(y) - f_1(x) |\,dy < \epsilon \qquad (0 \leqq x < x + h \leqq 1 + b).$$

Hence,

$$| F(x) - f(x) | < \epsilon \qquad (0 \leqq x < x + h \leqq 1). \tag{23}$$

Now apply (21) to $F(x)$, since $F'(x) = h^{-1}[f_1(x + h) - f_1(x)]$ is continuous in $(0, 1)$. This, combined with (23), yields[7]

$$| f(x) - F(0)x^{\nu-1} - \sum_{k=0}^{n} c_k x^{\nu_k} + c'_k x^{\bar{\nu}_k} |$$

$$\leqq | f(x) - F(x) | + | F(x) - F(0)x^{\nu-1} - \sum_{k=0}^{n} c_k x^{\nu_k} + c'_k x^{\bar{\nu}_k} | < \epsilon$$

$$(0 \leqq x \leqq 1).$$

THEOREM IV. *In order that the sequence* $\{l_k\}$ $(k = 1, 2, \cdots)$ *be a base for* $C(0, 1)$ *or* $L_2(0, 1)$, *it is necessary that condition* $(0\text{-}l_k)$ *be satisfied for any* $\psi(x) \in L_2(0, 1)$.

If

$$\int_0^1 y^{l_k}\psi(y)\,dy = 0 \qquad (k = 1, 2, \cdots),$$

[7] In the proof of (ii) the underlying idea is due to Szász, who applies (21) to x^l (l a positive integer) and then completes the proof by means of Weierstrass' Theorem. The introduction of Stekloff's function in this and similar problems is very helpful. Thus, it enables us to prove Weierstrass' Theorem for any continuous functions, once it is established for functions having a certain number of derivatives (by means of repeated integrals of the type (22)). The latter is readily achieved, e.g., by employing expansions in Fourier series.

then

$$\int_0^1 |\psi(y)|^2\,dy = \int_0^1 \psi(y)[\bar{\psi}(y) - c_n(y)]\,dy$$

$$\leqq \left[\int_0^1 |\psi(y)|^2\,dy \int_0^1 |\bar{\psi}(y) - c_n(y)|^2\,dy\right]^{\frac{1}{2}}, \qquad c_n(y) = \sum_{k=1}^{n} c_k y^{l_k},$$

(24) c_k = arbitrary constants.

$$\int_0^1 |\psi(y)|^2\,dy \leqq \left[\int_0^1 |\psi(y)|^2\,dy\right]^{\frac{1}{2}}$$

$$\cdot\left\{2\left[\int_0^1 |\bar{\psi}(y) - f(y)|^2\,dy + \int_0^1 |f(y) - c_n(y)|^2\,dy\right]\right\}^{\frac{1}{2}},$$

where $f(y)$, continuous in $(0, 1)$, is chosen so that

$$\int_0^1 |\bar{\psi}(y) - f(y)|^2\,dy < \epsilon.$$

Now, if $\{l_k\}$ is a base for $C(0, 1)$, by a proper choice of $c_n(x)$ we get

$$\int_0^1 |\psi(y)|^2\,dy < \left[\int_0^1 |\psi^2(y)|^2\,dy\right]^{\frac{1}{2}} \cdot \eta \qquad (\eta > 0, \text{ arbitrarily small}).$$

(If $\{l_k\}$ is a base for $L_2(0, 1)$, the same conclusion follows from (24).) If $\int_0^1 |\psi(y)|^2\,dy = 0$, our theorem is established. A contrary assumption leads to an absurd result; namely,

$$\int_0^1 |\psi(y)|^2\,dy < \eta^2.$$

THEOREM V. (i) *In order that* (18-C) *be a base for* $C(0, 1)$, *it is necessary that* $\sum_{k=1}^{\infty} (2\xi_k + 1)/[(\xi_k + 1)^2 + \eta_k^2]$ *diverge. This condition may be replaced by the divergence of* $\sum_{k=1}^{\infty} (\xi_k + 1)/(\xi_k^2 + \eta_k^2)$. (ii) *If we add the condition* $\xi_k \geqq \delta > 0$ *for* $k \geqq 1$, *the divergence of* $\sum_{k=1}^{\infty} (\xi_k + 1)/(\xi_k^2 + \eta_k^2)$ *is both necessary and sufficient.*

The first part follows directly, by combining Theorems III and IV and observing that the convergence of $\sum_{k=1}^{\infty} (\xi_k + 1)/(\xi_k^2 + \eta_k^2)$ implies that of $\sum_{k=1}^{\infty} (2\xi_k + 1)/[(\xi_k + 1)^2 + \eta_k^2]$. We pass to the second part (sufficiency).

Let $\alpha(x)$ be a real function, of bounded variations in $(0, 1)$, such that

$$\int_0^1 y^{\mu_k}\,d\alpha(y) = 0 \qquad (k = 0, 1, 2, \ldots) \tag{25}$$

$$(\mu_0 = 0; \quad \mu_k = \xi_k + i\eta_k; \quad \xi_k \geqq \delta > 0,\ k \geqq 1).$$

We may assume, without loss of generality, $\alpha(0) = 0$. (25), written for $k = 0$, yields $\alpha(1) = 0$. We get further, integrating by parts,

$$
(26) \qquad \begin{aligned} 0 &= y^{\mu_k}\alpha(y)\Big|_0^1 - \mu_k \int_0^1 y^{\mu_k-1}\alpha(y)\,dy, \\ &\int_0^1 y^{\mu_k-1}\alpha(y)\,dy = 0 \qquad (k = 1, 2, \cdots). \end{aligned}
$$

Letting

$$
y = e^{-lu}, \qquad l = \text{constant},\, l > \frac{1}{2} + \frac{1}{\delta},
$$

we transform (26) into

$$
(27) \qquad \int_0^1 e^{-t_k u}\phi(u)\,du = 0, \qquad t_k = l\mu_k, \qquad R(t_k) > \tfrac{1}{2} \qquad (k = 1, 2, \cdots).
$$

The function $\alpha(x)$ is bounded in $(0, 1)$, hence $\phi(u)$ in (27) is bounded in $(0, \infty)$. It follows that in (1), where $\phi(u)$ is the same as in (27), $F(1/(1 + z))$ is bounded in D. Observe further that the divergence of $\sum_{k=1}^{\infty} (\xi_k + 1)/(\xi_k^2 + \eta_k^2)$ implies here that of $\sum_{k=1}^{\infty} (2l\xi_k - 1)/[(l\xi_k)^2 + (l\eta_k)^2]$. By Theorem II, $\phi(u)$ necessarily vanishes almost everywhere in $(0, \infty)$ and so does $\alpha(x)$ in $(0, 1)$; in particular, $\alpha(x)$ vanishes at all points of continuity. Moreover, as we know, $\alpha(0) = \alpha(1) = 0$. Being of bounded variation, $\alpha(x)$ vanishes everywhere in $(0, 1)$, with the possible exception of a countable set of points, not containing 0, 1. Letting in (25)

$$
x^{\mu_k} = r_k(x) + is_k(x) \qquad (k = 0, 1, 2, \cdots),
$$

we get the equivalent relations

$$
(28) \qquad \int_0^1 r_k(y)\,d\alpha(y) = \int_0^1 s_k(y)\,d\alpha(y) = 0 \qquad (k = 0, 1, 2, \cdots).
$$

The relations (28), where $\alpha(y)$ has the aforesaid properties, show by virtue of a theorem of F. Riesz [10] that any function which is continuous in $(0, 1)$ may be approximated therein indefinitely by linear aggregates of the functions

$$
\{r_k(x) = \frac{1}{2}(x^{\mu_k} + x^{\bar{\mu}_k}),\, s_k(x) = \frac{1}{2i}\,(x^{\mu_k} - x^{\bar{\mu}_k})\} \qquad (k = 0, 1, 2, \cdots).
$$

Thus, the statement (ii) concerning (18-C) is established.

Remark. Since the two series $\sum_{k=1}^{\infty} (\xi_k + 1)/(\xi_k^2 + \eta_k^2)$, $\sum_{k=1}^{\infty} \xi_k/(\xi_k^2 + \eta_k^2)$ converge and diverge together under condition (ii), we conclude that *the divergence*

of $\sum_{k=1}^{\infty} \xi_k/(\xi_k^2 + \eta_k^2)$ *is necessary and sufficient in order that the sequence* (18-C), *satisfying the conditions of Theorem* V (ii), *be a base for* $C(0, 1)$.

As a particular case, we get the following results. *The infinite sequence* $\mu_0 = 0, \mu_1, \mu_2, \cdots, \mu_k, \cdots; \mu_k > 0, \varliminf_{k=\infty} \mu_k > 0$, *is a base for* $C(0, 1)$. *The sequence* $\mu_0 = 0 < \mu_1 < \mu_2 < \cdots$ *is a base for* $C(0, 1)$ *if and only if* $\sum_{k=1}^{\infty} \mu_k^{-1}$ *diverges* [12,7]. Combining the previous results with our lemma, we derive corresponding results for $L_2(0, 1)$.

THEOREM VI. *In order that* (18-L) *be a base for* $L_2(0, 1)$, *it is necessary that* $\sum_{k=0}^{\infty} (2\xi_k + 1)/[(\xi_k + 1)^2 + \eta_k^2]$ *diverge. If, in addition,* $\xi_k \geqq \delta > 0$ *for* $k \geqq 0$, *the divergence of* $\sum_{k=0}^{\infty} \xi_k/(\xi_k^2 + \eta_k^2)$ *is both necessary and sufficient* (*as for* (18-C)).[8]

4. We close our discussion with the following remarks.

(i) The above considerations can be readily extended to $L_p(0, 1)$, $p \geqq 1$, also to any sequence of continuous functions defined in (0, 1) taking the place of $\{x^{\mu_k}, x^{\bar{\mu}_k}\}$.

(ii) Writing

$$f(\mu) = \int_0^1 y^{\mu} \psi(y)\, dy = \int_0^1 y \cdot y^{\mu-1} \psi(y)\, dy, \tag{17}$$

we can discuss the representation (17) along the foregoing lines (uniqueness, approximation), making use of Jacobi polynomials $\phi_n(x; 0, 1; x)$. The results thus obtained may be applied to the Laplace integral (1), by means of the substitution $e^{-u} = y$.[9]

BIBLIOGRAPHY

1. L. BIEBERBACH, *Lehrbuch der Funktionentheorie*, Leipzig, 1927, vol. II.
2. W. FELLER, *Completely monotone sequences and functions*, this Journal, vol. 5(1939), pp. 661–674.
3. EINAR HILLE, *Bilinear formulas in the theory of the transformation of Laplace*, Compositio Mathematica, vol. 6(1938), pp. 93–102.
4. EINAR HILLE AND J. D. TAMARKIN, *On moment functions*, Proc. Nat. Acad. Sc., vol. 19(1933), pp. 902–908.
5. E. LANDAU, *Darstellung und Begründung einiger neuerer Ergebnisse der Funktionentheorie*, Berlin, 1916.
6. M. LERCH, *Sur un point de la théorie des fonctions génératrices d'Abel*, Acta Math., vol. 27(1903), pp. 339–352.

[8] Szász considers approximation, in the sense of (19-L), for continuous functions; his result is more complete. Some of our results (Theorems V and VI) differ non-essentially from the corresponding results of Szász.

[9] Cf. Widder [15], where the possible use of Jacobi polynomials is indicated for the purpose of inversion.

7. CH. H. MÜNTZ, *Ueber den Approximationssatz von Weierstrass*, Math. Abhandlungen H. A. Schwarz gewidmet, Berlin, 1914, pp. 303–312.
8. R. E. A. C. PALEY AND NORBERT WIENER, *Fourier Transforms in the Complex Domain*, American Mathematical Society Colloquium Publications, vol. 19, 1934.
9. M. PICONE, *Sulla transformazione di Laplace*, Rend. Atti Accad. Nat. Lincei, (6), vol. 21(1935), pp. 306–313.
10. F. RIESZ, *Sur certaines systèmes singulières des équations intégrales*, Ann. de l'Ecole Normale Sup., (3), vol. 28(1911), pp. 33–61.
11. W. STEKLOFF, *Sur la théorie de fermeture des systèmes de fonctions orthogonales dépendant d'un nombre quelconque de variables*, Mémoires (Zapiski) de l'Acad. Imp. des Sciences de St. Petersburg, (8), vol. 30(1911), pp. 1–86.
12. O. SZÁSZ, *Ueber die Approximation stetiger Funktionen durch lineare Aggregate von Potenzen*, Math. Ann., vol. 77(1915–1916), pp. 482–496.
13. E. C. TITCHMARSH, *The Theory of Functions*, Oxford, 1932.
14. F. TRICOMI, *Transformazione di Laplace e polinomi di Laguerre*, Rend. Atti Accad. Nat. Lincei, (6), vol. 21(1935), pp. 232–239.
15. D. V. WIDDER, *An application of Laguerre polynomials*, this Journal, vol. 1(1935), pp. 126–136.
16. A. WINTNER, *Bemerkung zum Eindeutigkeitssatz der Laplaceschen Transformierten*, Math. Zeits., vol. 36(1933), pp. 638–640.

UNIVERSITY OF PENNSYLVANIA.

AN AXIOMATIC CHARACTERIZATION OF L_p-SPACES

BY F. BOHNENBLUST

1. **Introduction.** The spaces

(a) $l_{p,n}$, $l_{p,\infty}$ $(1 \leqq p < \infty)$, whose elements are finite or infinite sequences $x = \{\xi_n\}$ with a finite norm $\| x \| = (\sum | \xi_n |^p)^{1/p}$,

(b) L_p $(1 \leqq p < \infty)$, whose elements are Lebesgue measurable functions over the interval (0, 1) with a finite norm $(\int | f(t) |^p \, dt)^{1/p}$,

(c) any direct sum of (a) and (b) with the same number p, where the norm is given as the $l_{p,2}$ norm of the norms of the components,

(d) $l_{\infty,n}$, whose elements are $\{\xi_1, \cdots, \xi_n\}$ with the norm max $(| \xi_n |)$,

(e) c_0, whose elements are sequences converging to zero with the norm max $(| \xi_n |)$,

are examples of separable, normed, complete, linear spaces (i.e., separable Banach spaces). They can be partially ordered by defining $x_1 < x_2$, if $\xi_n^{(1)} \leqq \xi_n^{(2)}$, for all n, when the space is a sequence space, $f_1(t) \leqq f_2(t)$ a.e. when the space is a function space. This partial ordering satisfies all axioms of §2. Finally, all these spaces have the following property:

PROPERTY P. *If $a = a_1 + a_2$, where a_1 and a_2 are orthogonal, if $b = b_1 + b_2$, where b_1 and b_2 are orthogonal, and if*

$$\| a_1 \| = \| b_1 \|, \qquad \| a_2 \| = \| b_2 \|,$$

then $\| a \| = \| b \|$.

The purpose of the present paper is to show that this property is characteristic for the spaces considered. Precisely, we prove the

THEOREM. *Any partially ordered, separable, Banach space of at least three dimensions in which property* P *is valid is strongly equivalent* (cf. §7 for terminology) *to one of the above-mentioned spaces.*

Some results are also obtained when the separability is not assumed.

2. **Partially ordered spaces.**[1]

2.1. **Axioms.** A linear space is said to be a partially ordered linear space if it satisfies the following axioms:

Received January 2, 1940; presented to the American Mathematical Society, December 2, 1939.

[1] Partially ordered spaces were considered first by F. Riesz (Proceedings of the International Congress of Mathematicians, Bologna, vol. 3, 1928, pp. 143–148) and by L. Kantorowitch (cf., in particular, *Lineare halb-geordnete Räume*, Recueil Mathématique, new series, vol. 2(1937), pp. 121–165). The present paper is based on H. Freudenthal's paper *Über teilweise geordnete Moduln*, Proceedings, Amsterdam Academy, vol. 39(1936), pp. 641–651. This paper will be cited as F. The axioms which we are postulating correspond to those in this paper up to and including (6.3).

AXIOM 1. *It is a lattice, in which every denumerable set, bounded from above, has a least upper bound, denoted by* $\vee S$.

AXIOM 2. *The order relation is invariant under translation and under multiplication by positive numbers, whereas it is reversed under multiplication by* -1.

AXIOM 3. *There exists an element, denoted by* 1, *such that* $1 \wedge a = 0$ *is true only for* $a = 0$.

The order relation $<$ of the lattice will be understood as an inclusion; i.e., it shall mean "less than or equal to". As Freudenthal and Kantorowitch have shown, a partially ordered linear space is a distributive lattice. It is distributive even with respect to denumerable sets, i.e.,

$$\bigvee_n a_n \wedge \bigvee_m b_m = \bigvee_{n,m} [a_n \wedge b_m],$$

and a similar relation holds with inf and sup interchanged.

2.2. **Notation.** With a few changes we adopt the notation of Freudenthal:

(i) The element $\lambda \cdot 1$ shall be denoted more simply by λ.

(ii) $a_+ = 0 \vee a$; $a_- = 0 \wedge a$; $|a| = a_+ - a_- = a \vee -a$.

(iii) Two elements a and b of the linear space are said to be orthogonal if, and only if, $|a| \wedge |b| = 0$.

(iv) E is the Boolean algebra of all elements of the linear space for which $0 < e < 1$ and $e \wedge (1 - e) = 0$.

(v) $e[a] = \bigvee_n [1 \wedge na_+]$.

(vi) $e(a; \xi) = e[a - \xi]$ or more simply $e(\xi)$, when no confusion need be feared as to the underlying element a.

We state one result of Freudenthal explicitly, partly because it will be of frequent use to us and partly because of a misprint in its statement as it appeared in Freudenthal's paper. For other results a reference, e.g., F(7.3), will be given.

F(7.5) *Let* γ_n *be a sequence of real numbers defined for all integers* n, *such that* $0 \leqq \gamma_{n+1} - \gamma_n \leqq \delta$; $\gamma_0 = 0$; $\lim \gamma_n = \pm\infty$ *as* n *tends to infinity. Then for any element* a

$$\begin{aligned}\bigvee_{n>0} \gamma_n[e(\gamma_n) - e(\gamma_{n+1})] &+ \bigwedge_{n<0} \gamma_n[e(\gamma_n) - e(\gamma_{n+1})] \\ &= \bigvee_{n>0} \gamma_n e(\gamma_n) + \bigwedge_{n<0} \gamma_n[1 - e(\gamma_{n+1})] < a \\ &< \bigvee_{n>0} \gamma_n e(\gamma_{n-1}) + \bigwedge_{n<0} \gamma_n[1 - e(\gamma_n)] \\ &= \bigvee_{n>0} \gamma_n[e(\gamma_{n-1}) - e(\gamma_n)] + \bigwedge_{n<0} \gamma_n[e(\gamma_{n-1}) - e(\gamma_n)].\end{aligned}$$

Defining stepelements to be elements s of the form $s = \bigvee \alpha_n e_n + \bigwedge \alpha_n e_n$, where the e_n belong to E and are mutually orthogonal, we see from F(7.5)

that corresponding to any element a and any real positive number δ there exists a stepelement s for which

$$s < a < s + \delta.$$

2.3. **The function** $e(\xi) = e(a; \xi)$. According to our notation (vi) each element a determines a function $e(\xi)$ defined over the real numbers with values in E. Different elements determine different functions. The elements e of E play the rôle of the planes of coördinates and the function $e(\xi)$ the rôle of the coördinates, or, more precisely, $e(\xi)$ is the plane of those coördinates which are greater than ξ. The main properties of the function $e(\xi)$, all of which are contained implicitly in the paper of Freudenthal, shall be briefly enumerated.

$$e(\xi) < e(\eta) \text{ for } \eta < \xi. \tag{2.3.1}$$

$$\bigwedge_{\xi} [e(\xi)] = 0; \qquad \bigvee_{\xi} [e(\xi)] = 1. \tag{2.3.2}$$

$$e(\xi) = \bigvee_{\eta > \xi} [e(\eta)]. \tag{2.3.3}$$

$$e(\alpha \cdot a; \xi) = e(a; \xi/\alpha) \text{ for } \alpha > 0. \tag{2.3.4}$$

$$e(-a; \xi) = \bigvee_{\eta > \xi} [1 - e(a; -\eta)] = 1 - \bigwedge_{\eta > \xi} e(a; -\eta). \tag{2.3.5}$$

$$e(a + b; \xi) = \bigvee_{\rho + \sigma = \xi} [e(a; \rho) \wedge e(b; \sigma)]. \tag{2.3.6}$$

The existence of the inf and the sup, even when they are to be formed over non-denumerable sets, can easily be established. In the last relation, in particular, ρ and σ can vary either over all real numbers or over any everywhere dense set. This relation is easily verified for stepelements and this result can be extended to the general case by making use of F(7.5).

We give a proof of (2.3.5), since this relation does not appear as such in Freudenthal's paper. Evidently we may assume $\xi = 0$. The relation takes then the form

$$e[-a] + \bigwedge_{\eta > 0} e[a + \eta] = 1.$$

Thus it is a slight improvement of F(7.3). From the definition of the element $e[a]$ it is immediate that the left member of this relation is > 1. On the other hand, from F(7.4.3) it follows that $e(-a; 1/m) < m(-a)_+$ and a fortiori

$$(1 - m(-a)_+) \wedge m(-a)_+ < 1 - e(-a; 1/m).$$

Since $e > 1 \wedge -a$ implies $e > (1 - m(-a)_+) \wedge m(-a)_+$, the alternate definition of $e[a]$ (see F(7.2)) shows that

$$(1 - m(-a)_+) \wedge m(-a)_+ < e[-a].$$

Combining these two inequalities, we obtain

$$(1 - m(-a)_+) \wedge m(-a)_+ < e[-a] - e(-a; 1/m),$$

and therefore for any positive integer n

$$1 \wedge n(-a)_+ \wedge n(1/m - (-a)_+) < e[-a] - e(-a; 1/m).$$

This relation implies by (2.3.3)

$$\begin{aligned} e[-a] \wedge [\bigwedge_m e[a + 1/m]] \\ = e[(-a)_+] \wedge [\bigwedge_m e[1/m - (-a)_+]] < e[-a] - \bigvee_m e(-a; 1/m) = 0, \end{aligned}$$

and equation (2.3.5) follows from it readily.

3. **Normed partially ordered linear spaces.** *P***-spaces.** A normed, partially ordered, linear space shall be a space as considered in the previous section in which, in addition, a norm is assumed, with respect to which the space is complete. We do not assume any of the axioms which Kantorowitch or Freudenthal postulate and which correlate the norm and the partial ordering. On the other hand, we assume the following

PROPERTY P. *If a and b are two elements of the linear space which can be split up as the sum of two mutually orthogonal elements*:

$$a = a_1 + a_2, \; |a_1| \wedge |a_2| = 0; \qquad b = b_1 + b_2, \; |b_1| \wedge |b_2| = 0,$$

in such a way that $\|a_1\| = \|b_1\|$ and $\|a_2\| = \|b_2\|$, then the norms of the elements a and b must be equal.

For simplicity we shall call *any normed partially ordered linear space in which property* P *is valid a P-space.*

In §8 of F it is verified that any element e of the Boolean algebra E induces a decomposition of the space as the direct sum of two subspaces: $F = F_e \oplus F_{1-e}$. Any two elements a and b, a in F_e, b in F_{1-e}, are orthogonal to each other, and if the space is a P-space, the projection P_e induced by the decomposition will have a norm equal to one (assuming of course $e \neq 0$). In particular,

THEOREM 3.1. *If the elements a_k tend to a according to the norm, their projections $P_e(a_k)$ will tend to $P_e(a)$.*

Finally we mention a theorem, which follows immediately from the results of §8 in F.

THEOREM 3.2. *If e_n is a sequence of mutually orthogonal elements of E such that $\bigvee e_n = 1$, then for any element $a > 0$*

$$a = \bigvee P_{e_n}(a).$$

4. **A functional equation.** Let f be any function defined over the quadrant ξ, $\eta \geqq 0$ of the (ξ, η)-plane, which satisfies the following conditions:

$$(4.1) \qquad f(\tau\xi, \tau\eta) = \tau f(\xi, \eta) \quad \text{for any } \xi, \eta, \tau \geqq 0,$$

$$f(\xi, \eta) \leqq f(\xi', \eta') \quad \text{for any } \xi' \geqq \xi \geqq 0, \eta' \geqq \eta \geqq 0, \tag{4.2}$$

$$f(\xi, \eta) = f(\eta, \xi) \quad \text{for any } \xi, \eta \geqq 0, \tag{4.3}$$

$$f(0, 1) = 1, \tag{4.4}$$

and finally the functional relation

$$f(\xi, f(\eta, \zeta)) = f(f(\xi, \eta), \zeta) \quad \text{for any } \xi, \eta, \zeta \geqq 0. \tag{4.5}$$

Any such function generates a sequence $\{\alpha_n\}$ $(n = 1, 2, \cdots)$ defined inductively by

$$\alpha_1 = 1; \qquad \alpha_n = f(1, \alpha_{n-1}) \quad \text{for } n > 1.$$

This sequence is evidently non-decreasing. Furthermore we have

LEMMA 4.1. *$\alpha_{n+m} = f(\alpha_n, \alpha_m)$ for any integers $n, m > 0$.*

This is equivalent to $\alpha_n = f(\alpha_m, \alpha_{n-m})$ for any $n > 1$ and any $m = 1, 2, \cdots, n - 1$. For $m = 1$ the relation holds by definition, and the general case can be deduced by induction:

$$\alpha_n = f(\alpha_{m-1}, \alpha_{n-m+1}) = f(\alpha_{m-1}, f(1, \alpha_{n-m}))$$
$$= f(f(1, \alpha_{m-1}), \alpha_{n-m}) = f(\alpha_m, \alpha_{n-m}).$$

LEMMA 4.2. *$\alpha_n \cdot \alpha_m = \alpha_{nm}$ for any integers $n, m > 0$.*

The proof is again by induction, n being fixed and m varying from one to infinity. The lemma holds trivially for $m = 1$. Assuming it to be true for $m - 1$, we get

$$\alpha_n \cdot \alpha_m = \alpha_n \cdot f(1, \alpha_{m-1}) = f(\alpha_n, \alpha_{nm-n}) = \alpha_{nm}.$$

Lemma 4.2 is almost Hammel's functional equation, but we are dealing with sequences rather than functions over real numbers. The monotone character of α_n implies $\alpha_2 \geqq 1$. Two cases can thus arise:

Case 1: $\alpha_2 = 1$; and

Case 2: $\alpha_2 > 1$.

The first case can be disposed of immediately. A repeated application of Lemma 4.2 shows that $\alpha_{2^n} = 1$ and thus that every $\alpha_n = 1$. Furthermore for any ξ $(0 \leqq \xi \leqq 1)$

$$1 = f(1, 0) \leqq f(1, \xi) \leqq f(1, 1) = 1,$$

so that $f(1, \xi) = 1$. Finally, by making use of condition (4.1),

$$f(\xi, \eta) = \max(\xi, \eta) = (\xi^p + \eta^p)^{1/p} \quad \text{for} \quad p = \infty.$$

Let us consider now Case 2: $\alpha_n > 1$ for $n > 1$.

LEMMA 4.3. *$\alpha_n = n^{1/p}$ for some real number $p > 0$.*

Proof. Let m and n be two fixed integers > 1. For any integer $k > 0$ determine the integer h for which

$$m^h \leqq n^k < m^{h+1}.$$

By Lemma 4.2 $\alpha_m^h = \alpha_{m^h}$, and similarly for k and $h + 1$. Hence

$$h \log \alpha_m \leqq k \log \alpha_n \leqq (h + 1) \log \alpha_m .$$

Substituting for h its expression in terms of k, dividing by k and letting k tend to infinity, we obtain

$$\frac{\log \alpha_m}{\log m} = \frac{\log \alpha_n}{\log n}.$$

In other words $\log \alpha_n/\log n$ is independent of n. In particular, if we put $\log \alpha_2/\log 2$ equal to $1/p$, we obtain the desired result.

LEMMA 4.4. *$f(1, r^{1/p}) = (1 + r)^{1/p}$ for any rational number $r > 0$.*

Proof. Put $r = m/n$, then

$$\begin{aligned} f(1, m^{1/p}/n^{1/p}) &= n^{-1/p}f(n^{1/p}, m^{1/p}) = n^{-1/p}f(\alpha_n , \alpha_m) \\ &= n^{-1/p}(n + m)^{1/p} = (1 + r)^{1/p}. \end{aligned}$$

It is easy to extend Lemma 4.4 to all real numbers and finally to prove, in combination with Case 1:

THEOREM 4.1. *Any function f satisfying the five conditions* (4.1) *to* (4.5) *is necessarily of the form*

$$f(\xi, \eta) = (\xi^p + \eta^p)^{1/p}$$

for some real number p $(0 < p \leqq \infty)$.

5. **Application to P-spaces.** Let us assume that we are dealing with a P-space which is at least three-dimensional. For any two elements $a, b > 0$ such that $a \wedge b = 0$, a function $f_{a,b}$ of two variables $\xi, \eta \geqq 0$ can be defined by

$$f_{a,b}(\xi, \eta) = \left\| \xi \frac{a}{\| a \|} + \eta \frac{b}{\| b \|} \right\|.$$

The property P shows that this function does not depend on the choice of the elements a and b. Accordingly, the indices a, b shall be dropped and the function will be denoted simply by f. Although f is defined for all values of ξ, η, we shall always restrict ourselves to non-negative values. For any $\tau > 0$:

$$f(\tau\xi, \tau\eta) = \tau f(\xi, \eta). \tag{5.1}$$

$$f(\xi, \eta) \leqq f(\xi', \eta') \quad \text{whenever} \quad \xi \leqq \xi' \quad \text{and} \quad \eta \leqq \eta'. \tag{5.2}$$

Assume first $\xi = \xi'$. In this case the relation (5.2) is true, because the element $\xi \cdot a/\| a \| + \eta \cdot b/\| b \|$ lies on the segment of endpoints $\xi \cdot a/\| a \| + \eta' \cdot b/\| b \|$

and $\xi \cdot a/\| a \| - \eta' \cdot b/\| b \|$. These endpoints have by property P the same norm. Similarly if $\eta = \eta'$. Then

$$f(\xi, \eta) \leqq f(\xi, \eta') \leqq f(\xi', \eta').$$

$$f(\xi, \eta) = f_{a,b}(\xi, \eta) = f_{b,a}(\eta, \xi) = f(\eta, \xi). \tag{5.3}$$

It is readily shown that the more general property P, referring to sums of any finite number of elements instead of only two, is a consequence of this special case. Accordingly, if $a_i > 0$ $(i = 1, \cdots, m)$ are mutually orthogonal, the function f of m variables ξ_i

$$f(\xi_1, \cdots, \xi_m) = \left\| \sum \xi_i \frac{a_i}{\| a_i \|} \right\|$$

does not depend on the choice of the a_i and in particular the function f is symmetric in the ξ_i. Since the P-space has been assumed to be at least three-dimensional, we can take $m = 3$. For $\xi_i > 0$ we have

$$\begin{aligned} f(\xi_1, f(\xi_2, \xi_3)) &= \left\| \frac{\xi_1 \cdot a_1}{\| a_1 \|} + f_{a_2, a_3}(\xi_2, \xi_3) \cdot \frac{\xi_2 a_2/\| a_2 \| + \xi_3 a_3/\| a_3 \|}{\| \xi_2 a_2/\| a_2 \| + \xi_3 a_3/\| a_3 \| \|} \right\| \\ &= \left\| \sum \xi_i \frac{a_i}{\| a_i \|} \right\| = f(\xi_1, \xi_2, \xi_3). \end{aligned}$$

The symmetry of f then implies

$$f(f(\xi_1, \xi_2), \xi_3) = f(\xi_1, f(\xi_2, \xi_3)). \tag{5.4}$$

Finally

$$f(1, 0) = 1. \tag{5.5}$$

The function f is of the type considered in §4. Consequently for some p $(0 < p \leqq \infty)$,

$$f(\xi, \eta) = (\xi^p + \eta^p)^{1/p}.$$

The convexity of the norm requires of course $p \geqq 1$. By induction we prove

THEOREM 5.1. *Corresponding to any P-space of dimension* $\geqq 3$ *there exist*[8] *a real number* p $(1 \leqq p \leqq \infty)$ *such that for any mutually orthogonal* a_i $(i = 1, \cdots, m)$

$$\| \sum a_i \| = (\sum \| a_i \|^p)^{1/p}.$$

6. **Generalized L_p-spaces. p finite.** As Freudenthal remarked, to any Boolean algebra E, where the inf and the sup exist for denumerable sets, there exists a partially ordered linear space whose Boolean algebra is the given one. We find it necessary to consider the construction of such a space in more details.

Let E be given. An element of the linear space F shall be any function $e(\xi)$, defined over the real numbers with values in the Boolean algebra E and which

satisfies the conditions (2.3.1), (2.3.2), and (2.3.3) of §2. Addition and multiplication by real numbers are defined by (2.3.4), (2.3.5) and (2.3.6). With these operations the functions $e(\xi)$ form a linear space F. The elements e of the Boolean algebra can be imbedded in F:

$$e \to e(\xi) = \begin{cases} 1 & \text{when } \xi < 0, \\ e & \text{when } 0 \leqq \xi < 1, \\ 0 & \text{when } 1 \leqq \xi. \end{cases}$$

The zero of E becomes the zero of F.

If we define an order in F by

$$e_1(\xi) < e_2(\xi) \textit{ if and only if } e_1(\xi) < e_2(\xi) \textit{ for every } \xi,$$

the space F is partially ordered. The unit element of E can be used as the 1 of the space F.

Let us consider now a proper measure or dimension function[2] $\mu(e)$ in E; i.e., a function $\mu(e)$, real valued, defined over E and satisfying

(6.1) $\mu(e) \geqq 0, = 0$ *if and only if* $e = 0$,

(6.2) $\sum \mu(e_n) = \mu(\vee e_n)$ *for any sequence* e_n *of mutually orthogonal elements of* E.

Definition 6.1. A positive element a of F, given by $e(\xi)$, will be said to be *p-integrable* ($1 \leqq p < \infty$) if there exists a sequence γ_n, $\gamma_0 = 0$, $0 \leqq \gamma_{n+1} - \gamma_n \leqq \delta$, $\lim \gamma_n = +\infty$, as n tends to infinity, such that

$$\sum_{n=1}^{\infty} \gamma_n^p \mu(e(\gamma_n) \wedge [1 - e(\gamma_{n+1})]) < \infty.$$

Definition 6.2. An element a of F is p-integrable if the element $|a|$ is p-integrable.

Evidently every element of E is p-integrable; furthermore, if a and b are p-integrable, then also $\alpha \cdot a$, $a \vee b$, $a \wedge b$, $a + b$ are p-integrable. Finally $\vee a_n$, $\wedge a_n$ are p-integrable if the sequence a_n is bounded by a p-integrable element of F. It is hardly necessary to verify these statements in detail.

Theorem 6.1. *The p-integrable elements form a partially ordered linear space whose Boolean algebra is E.*

There are no difficulties in introducing an integral starting from the approaching sums of Definition 6.1. With such an integral we introduce a norm for the space of p-integrable elements.

Definition 6.3. The norm $\|a\|_p$ of $a > 0$ is defined as the

$$\lim \left(\sum \gamma_n^p \mu(e(\gamma_n) \wedge [1 - e(\gamma_{n+1})])\right)^{1/p}$$

[2] Cf., e.g., J. v. Neumann, *Continuous geometries*, Proc. Nat. Acad. Sci., vol. 22(1936), pp. 92–100.

as the mesh δ of the subdivision tends to zero. For any element a

$$\| a \|_p = \| \, | a | \, \|_p .$$

The proof that $\| a \|_p$ exists and has all the properties of a norm follows too closely the classical methods used for Lebesgue L_p-spaces to be repeated here. The completeness of the space can also be shown, the limiting element of a Cauchy sequence is given, of course, by the o-convergence of Kantorowitch.[3]

DEFINITION 6.4. Any such normed partially ordered linear space shall be called a generalized L_p-space.

Property P is valid in any generalized L_p-space; in other words we have

THEOREM 6.4. *Any generalized L_p-space is a P-space.*

The converse of this theorem is true when the real number corresponding to the P-space is finite. First of all we remark that in L_p the norm of a stepelement is given by

$$\| s \|_p = \| \bigvee \alpha_n e_n + \bigwedge \alpha_n e_n \|_p = [\sum | \alpha_n |^p \mu(e_n)]^{1/p}.$$

Assume now that F is a given P-space with the norm $\| a \|$.

LEMMA 6.1. *Suppose $\alpha_n \geqq 0$, e_n are mutually orthogonal in E, and $\bigvee e_n = 1$. Assume also that either*

(A) $\alpha_n e_n < b$, b in F, or

(B) $\sum \alpha_n^p \| e_n \|^p < \infty$.

Put $s_N = \bigvee_1^N \alpha_n \cdot e_n$, then $s = \bigvee \alpha_n e_n$ exists and s_N tends to s according to the norm in F.

Condition (A) implies condition (B). For if (A) is valid, s exists and the series in (B) is bounded by $\| s \|^p$. It suffices thus to verify the lemma assuming condition (B) to hold. In this case the sequence s_n is a Cauchy sequence and there will exist a limiting element t. By Theorem 3.1, for any element e of E, $P_e(s_n)$ tends to $P_e(t)$. In particular

$$P_{e_m}(t) = \lim P_{e_m}(s_n) = \alpha_m e_m ,$$

and by Theorem 3.2 $t = \bigvee \alpha_n \cdot e_n$.

If the $\alpha_n = 1$, condition (A) is valid for $b = 1$ and we obtain

COROLLARY 6.1. *The function $\mu(e) = \| e \|^p$ is a proper measure function in E.*

The corresponding space L_p can be introduced.

COROLLARY 6.2. *The stepelements belonging to F belong to L_p, and conversely. Their norms in F and in L_p are equal.*

It is only a restatement of the equivalence of conditions (A) and (B) in the preceding lemma.

[3] (Added March 6, 1940.) Such spaces have been considered more exhaustively and under more general assumptions by F. Wecken (Math. Zeit., vol. 45(1939), pp. 377–404) and by J. Olmsted in an article as yet unpublished.

LEMMA 6.2. *Every element in F belongs to L_p and* $\| a \|_p \leqq \| a \|$.

By F(7.5) there exist stepfunctions such that

$$s_1 < | a | < s_2 .$$

LEMMA 6.3. *Every element in L_p belongs to F.*

Let a be any element in L_p. The stepelements are everywhere dense in L_p. There exists a sequence s_n which tends to a according to the norm in L_p. The s_n form a Cauchy sequence in L_p and by Lemma 6.2 a Cauchy sequence in F. There exists an element b in F which is the limit of s_n in F. By Lemma 6.3 this element b is in L_p and

$$\| b - s_n \|_p \leqq \| b - s_n \|.$$

Thus the s_n tend to b in L_p and b must be equal to a. The proof shows also that the stepelements are everywhere dense in F and finally

THEOREM 6.5. *Any P-space whose characteristic number p is finite is an L_p-space, provided it is at least three-dimensional.*

7. **Special cases.** Two Banach spaces are said to be equivalent if there exists a norm preserving one-to-one correspondence between their elements. If the spaces are partially ordered, and the correspondence can be chosen to be also order preserving, they will be said to be strongly equivalent. Naturally we do not assume that the unit elements are transformed into each other under the correspondence. If the Boolean algebra is atomistic (i.e., if corresponding to any element e of E, different from the zero element, there always exists an atom contained in it) there will be a countable number of atoms, and it is easily seen that the P-space is strongly equivalent to a finite dimensional l_p-space (ordered according to the coördinates) or to the infinite dimensional one.

It is possible to decompose our Boolean algebra into a sub-Boolean algebra containing all the atoms and a complementary one free of atoms. If the P-space is separable, this complementary algebra will also be separable and following Freudenthal's method[4] for Hilbert space, we see that the corresponding L_p-space is strongly equivalent to the classical L_p-space on the interval (0, 1). We obtain thus

THEOREM 7.1. *Any separable P-space of dimension* $\geqq 3$ *whose characteristic number p is finite is strongly equivalent to either*

(i) *the finite dimensional* $l_{p,n}$,
(ii) *the infinite dimensional* l_p,
(iii) *the space* L_p *over* (0, 1), *or*
(iv) *the direct sum of* (i) *and* (iii), *or* (ii) *and* (iii).

[4] H. Freudenthal, loc. cit., pp. 650–651.

8. **P-spaces with $p = \infty$.** For any finite number of mutually orthogonal elements $a_1, \cdots, a_n$ of a P-space with $p = \infty$, the norm of their sum is the maximum of the norms of the terms. This relation can be applied to the sums

$$e_1 \vee e_2 = e_1 + (e_2 \wedge e_1') = e_2 + (e_1 \wedge e_2'),$$

where e_1 and e_2 are any two elements of the Boolean algebra of the P-space. If $e_1 < e_2$, we obtain in particular $\| e_1 \| \leqq \| e_2 \|$. In the general case:

$$\begin{aligned} \| e_1 \vee e_2 \| &= \max [\| e_1 \|, \| e_2 \wedge e_1' \|, \| e_2 \|, \| e_1 \wedge e_2' \|] \\ &= \max [\| e_1 \|, \| e_2 \|]. \end{aligned}$$

By induction, this equation is easily generalized to any finite number of e's. It remains true even for an infinite sequence e_n, if $\| e_n \| \to 0$ as $n \to \infty$; for in such a case the partial $\vee$'s form a Cauchy sequence. In other words, the norm of the P-space determines in E a measure function $\mu(e) = \| e \|$ with the properties:

(8.1) $\mu(e) \geqq 0$; $\mu(e) = 0$ *if, and only if*, $e = 0$,

(8.2) $\mu(e_n) \to 0$ *as* $n \to \infty$ *implies* $\mu(\vee e_n) = \max \mu(e_n)$.

These two properties will turn out to be characteristic. But before discussing this point, we investigate the further properties of E under the additional assumption that the P-space is separable. Any subset of a metric, separable space is separable, thus the Boolean algebra E will also be separable.

THEOREM 8.1. *A Boolean algebra is necessarily atomistic if it has the following properties*:

(i) $\vee$, $\wedge$ *exist for denumerable sets.*

(ii) *There exists a measure function* $\mu(e)$, *satisfying* (8.1) *and* (8.2).

(iii) *The Boolean algebra is separable* (*with respect to the topology of the metric* $d(e, f) = \mu[(e \wedge f') \vee (e' \wedge f)]$ *induced by* (ii).

We remark that it will be sufficient to verify that the Boolean algebra possesses atoms. For, the subalgebra of all elements less than a given element will satisfy the same conditions of the theorem.

LEMMA 8.1. *If* α *is any element of a certain set* A, *and if for any* α *there corresponds an element* e_α *of* E *such that*

(i) $e_\alpha \neq 0$,

(ii) $e_\alpha \wedge e_\beta = 0$ *for* $\alpha \neq \beta$,

(iii) $\mu(e_\alpha) \geqq \delta > 0$,

then the set A *is finite.*

Let A be infinite. Its cardinal number is at least $\aleph_0$. There will exist thus a sequence e_n with the properties (i), (ii), (iii). For any function f defined over all positive integers and whose values are either 0 or 1, we form the element

$$e_f = \vee\, e_n^{f(n)}, \quad \text{where} \quad e_n^0 = 0, \qquad e_n^1 = e_n.$$

If f and g are two different functions, they will differ for a certain integer m, say $f(m) = 0$, whereas $g(m) = 1$. It is readily shown that the element e_m is $< e_g \wedge e_f'$, and thus that the distance between the elements e_f and e_g is at least δ. Let now x_n be everywhere dense in E. Corresponding to each f we construct the set X_f consisting of all elements x_n which are at a distance less than $\frac{1}{2}\delta$ from e_f. No such set is void and they are disjoint. The sets X_f and thus the functions f could be enumerated, and this is a contradiction.

Lemma 8.1 obviously implies

LEMMA 8.2. *If for α in Λ*
(i) $e_\alpha \neq 0$,
(ii) $e_\alpha \wedge e_\beta = 0$ *for* $\alpha \neq \beta$,
then the set Λ is denumerable and for any enumeration $\mu(e_n) \to 0$ as $n \to \infty$.

Making use of the condition (8.2) imposed on the measure function, we obtain

LEMMA 8.3. *If e_α is any class of mutually orthogonal elements of E, the $\vee\, e_\alpha$ exists and $\mu(\vee\, e_\alpha) = \max \mu(e_\alpha)$.*

LEMMA 8.4. *If an element e of E is such that $0 \neq f < e$ implies $\mu(f) = \mu(e)$, then the Boolean algebra determined by e is finite.*

Proof. It is sufficient to verify this lemma for $e = 1$ and $\mu(e) = 1$, i.e., to verify that the Boolean algebra is finite if the measure of every one of its elements is 1, with the exception of the zero element. We prove first that E has atoms. Assume there are no atoms. Let e_1 be any element $\neq 0$ and $\neq 1$ and choose e_n inductively to satisfy the conditions
(i) $e_n < (e_1 \vee e_2 \vee \cdots \vee e_{n-1})'$,
(ii) $e_n \neq 0$ and $\neq (e_1 \vee e_2 \vee \cdots \vee e_{n-1})'$.
We would obtain an infinite sequence of mutually orthogonal elements of measure 1 against Lemma 8.1. Every subalgebra will have atoms; in other words, E is atomistic. Finally Lemma 8.1 again guarantees that there can be only a finite number of atoms.

LEMMA 8.6. *If there exists in E an element e and a positive number δ such that for any element $f < e$ either*

$$\mu(f) \leqq \mu(e) - \delta \quad or \quad \mu(f' \wedge e) \leqq \mu(e) - \delta,$$

then there exists an atom $< e$.

Again we may assume that the element e is the unit element 1, and let us assume $\mu(1) = 1$. Let e_α be a well-ordering of all elements of E. Define f_α by transfinite induction as follows:

$$f_1 = \begin{cases} e_1 & \text{if } \mu(e_1) \leqq 1 - \delta, \\ 0 & \text{otherwise.} \end{cases}$$

If for a certain β all f_α with $\alpha < \beta$ are already defined, put

$$f_\beta = \begin{cases} e_\beta & \text{if } e_\beta \wedge f_\alpha = 0 \text{ for every } \alpha < \beta \text{ and } \mu(e) \leqq 1 - \delta, \\ 0 & \text{otherwise.} \end{cases}$$

Any two f_α corresponding to different indices are orthogonal. By Lemma 8.3 $\vee f_\alpha$ exists (let us call it f), and

$$\mu(f) = \max \mu(f_\alpha) \leqq 1 - \delta.$$

We see in particular that $f \neq 1$ and therefore that $f' \neq 0$. The Boolean algebra determined by f' contains at least two elements. Let g be any one element of it, different from 0. For a certain α_1 we must have $g = e_{\alpha_1}$. Since f' is orthogonal to every f_α, g is a fortiori orthogonal to every f_α. In particular, it is orthogonal to f_{α_1}, and since $g \neq 0$, f_{α_1} must be the zero element. From the definition of f_{α_1} we see also that $\mu(g) > 1 - \delta$. From our assumption in the lemma follows $\mu(g') \leqq 1 - \delta$ and finally from property (8.2) of the measure function follows $\mu(g) = 1$. In other words, the Boolean algebra determined by f' satisfies the conditions of Lemma 8.3. It is therefore finite, and f' being different from 0, it must possess atoms.

It is now very easy to verify Theorem 8.1. Let us assume that E has no atoms. Then by the last lemma, corresponding to any element e and any $\epsilon > 0$, there exists e_1 and e_2, $e_2 \wedge e_1 = 0$, $e_1 \vee e_2 < e$ and such that $\mu(e_1) \geqq \mu(e) - \epsilon$, $\mu(e_2) \geqq \mu(e) - \epsilon$. We can determine inductively a sequence e_n as follows:
e_1 : any element for which $\mu(e_1) > \frac{1}{2}$, $\mu(e_1') = 1$;
e_n: $e_n < (e_1 e_2 \cdots e_{n-1})'$, $\mu(e_n) > \frac{1}{2}$, and $\mu([e_1 e_2 \cdots e_n]') = 1$.
This existence of such a sequence is in contradiction to the result of Lemma 8.1. Lemma 8.2 gives us furthermore

THEOREM 8.2. *The number of atoms in E is at most denumerable.*

We return to the consideration of the separable P-space. Its Boolean algebra is atomistic, with atoms $e_1, \cdots, e_n, \cdots$, either finite in number or denumerable. Every element of the space is a stepelement, with the coördinates α_n. If there are only a finite number of atoms, the norm of any element will be given by

$$\| a \| = \max [| \alpha_n | \, \| e_n \|];$$

and the P-space is strongly equivalent to $l_{\infty,n}$. Next we consider the case where there exists an infinite number of atoms. The same argument with which we proved Lemma 8.1 shows that in this case $| \alpha |_n \cdot \| e_n \|$ must tend to zero as n tends to infinity. Conversely if the α_n are chosen arbitrarily, but so as to satisfy $\alpha_n \| e_n \| \to 0$, there will exist an element of the P-space, whose coördinates are the chosen α_n. This is true, for the partial sums will form a Cauchy sequence. This fact allows us also to state that the norm is given by the same expression as in the finite case. We formulate these results in

THEOREM 8.3. *If the P-space is separable, and if its characteristic number p is infinite, then the space is strongly equivalent to either*

(i) *the finite dimensional $l_{\infty,n}$, or to*

(ii) *the space of sequences converging to zero, with the maximum of the absolute values of the coördinates for the norm.*

9. **General *P*-spaces, $p = \infty$.** The separability is no longer postulated in this section.

THEOREM 9.1. *Let E be a Boolean algebra, for which* $\vee$, $\wedge$ *exist for bounded denumerable sets, and in which a measure function* $\mu(e)$, *satisfying* (8.1) *and* (8.2), *is defined. There exists a P-space, with* $p = \infty$ *and with Boolean algebra E, such that in E the norm and the measure function coincide.*

The construction of the P-space is not unique. We give one example. An element of the space is any function $e(\xi)$ satisfying (2.3.1) to (2.3.3) for which

$$\xi \cdot \mu(e(\xi)) \qquad \text{and} \qquad -\xi \cdot \mu(1 - e(\xi))$$

are bounded. The least upper bound is taken to be the norm of the element. Addition and multiplication by real numbers are again defined by (2.3.4) to (2.3.6). The space thus obtained will satisfy all conditions required in Theorem 9.1. The verification of this fact is straightforward and will be omitted.

The space which we constructed corresponds, evidently, to the space of bounded sequences, and accordingly it will be denoted by $m(E, \mu)$. Other spaces could have been constructed, in particular, the space corresponding to sequences converging to zero. In a certain sense, however, the spaces $m(E, \mu)$ are universal.

THEOREM 9.2. *Let F be any P-space with* $p = \infty$ *and let* $m(E, \mu)$ *be the space constructed over the Boolean algebra of F and the measure function given by the norm of F. The functions* $e(\xi)$ *determined by the elements of F all belong to* $m(E, \mu)$.

We may rephrase this:

Any P-space is a linear subspace of its corresponding $m(E, \mu)$.

It should be remarked, however, that this imbedding in $m(E, \mu)$ is unsatisfactory in the sense that the metrics in F and in $m(E, \mu)$ may not coincide or may even be incomparable.

F(7.5) shows that we may restrict our proof to positive stepelements $s = \vee\, \alpha_n e_n$, where the α_n tend to infinity with n. Property P, for $p = \infty$, implies $\alpha_n \| e_n \| \leqq \| s \|$, and hence $\| e_n \| \to 0$ as $n \to \infty$. Finally, for $\xi > 0$,

$$\xi \| e(\xi) \| = \xi \| \bigvee_{\alpha_n > \xi} e_n \| = \max_{\alpha_n > \xi} \xi \| e_n \| \leqq \text{l.u.b. } \alpha_n \| e_n \| \leqq \| s \|.$$

A similar proof holds for negative stepelements, and the result can be extended to any element.

PRINCETON UNIVERSITY AND CALIFORNIA INSTITUTE OF TECHNOLOGY.

RESIDUATED DISTRIBUTIVE LATTICES

By Morgan Ward

I. Introduction

1. Although the lattice analogues of E. Noether's decomposition theorems for the ideals of commutative rings are independent of the modular or distributive condition,[1] it is of interest to consider the decompositions when such conditions are imposed. We study here the frequently occurring case when the basic lattice is distributive.[2]

Let $\mathfrak{S}$ be a residuated lattice in which the ascending chain axiom holds. If every irreducible of $\mathfrak{S}$ is primary, $\mathfrak{S}$ is called a "Noether lattice"; for such lattices, all the decomposition theorems of general commutative ideal theory are valid. For brevity, call such a lattice an N-lattice.

If $a \supset b$ in $\mathfrak{S}$ if and only if there exists an element c such that $ac = b$, $\mathfrak{S}$ is called a "principal element lattice", or for short a P-lattice. A P-lattice is distributive (W-D [1], Theorem 13.2) and so is every lattice in which any one of the following three equivalent conditions holds (W-D [1], Theorem 13.1):

$$(1.1) \qquad a:b \cup b:a = i; \qquad a:(b \cap c) = a:b \cup a:c; \qquad (b \cup c):a = b:a \cup c:a.$$

Here i is the identity element of the lattice.

A P-lattice is necessarily an N-lattice (W-D [1], Theorem 13.3). There exist, however, finite residuated lattices in which conditions (1.1) hold but which are not N-lattices. A Noether lattice in which conditions (1.1) hold will be called a "semi-arithmetical lattice", or SA-lattice.

2. Our main results are as follows. Necessary and sufficient conditions that an N-lattice be a P-lattice are, first, that an element be irreducible if and only if it is a power of a prime; and secondly, if p is a prime and q any proper divisor of p, then $qp = p$. A necessary and sufficient condition that an N-lattice be an SA-lattice is that any two primaries of the lattice be either coprime or else divisible one by the other.

In both types of lattice, the Noether decomposition of an element as a cross-cut of primary components is unique; this unicity is not a property of an arbitrary distributive Noether lattice as we show by an example, contrary to a statement in W-D [1].

In both an SA-lattice and a P-lattice, every element has a unique reduced

Received January 29, 1940.

[1] See the paper Ward-Dilworth [1] cited at the close of this article. We shall refer to this paper as W-D [1]. Numbers in brackets refer to the list of references at the end.

[2] The previous investigations in Ward [1], [2], appear here as special cases.

multiplicative decomposition into a product of primaries. A Noether lattice which is both an SA-lattice and a P-lattice is an arithmetical lattice, that is, a direct cross-cut of simple chains (see W-D [2] or W [1]). But not every arithmetical lattice is a P-lattice, nor is every residuated lattice with ascending chain condition which is a direct product of simple chains a Noether lattice.

A simple instance of a P-lattice is a commutative ring in which every ideal is principal; if the ring is also a domain of integrity, it is an SA-lattice as well.

The plan of the paper is indicated by the chapter headings. We assume the reader is familiar with the notation and terminology of W-D [1]; we shall adhere to this save that we employ $x \cup y$ and $x \cap y$ instead of (x, y) and $[x, y]$ for union and cross-cut, and we call divisor-free elements "simple".

II. Principal element lattices

3. The following three lemmas are valid in any residuated lattice with ascending chain condition, and may be proved as in the ideal theory of commutative rings.

LEMMA 3.1. *If q is a primary element of $\mathfrak{S}$, there exists a prime p such that*

$$p \supset q \supset p^k, \qquad q \not\supset p^{k-1}. \tag{3.1}$$

k here is a positive integer. p is called the prime corresponding to q, and is uniquely determined by (3.1).

LEMMA 3.2. *If q is a primary element of $\mathfrak{S}$, and a is any element such that $q \not\supset a$, then $q{:}a$ is primary and corresponds to the same prime as q.*

LEMMA 3.3. *If q is a primary element of $\mathfrak{S}$ and p is the prime corresponding to q, and a is any other element of $\mathfrak{S}$, then*

$$q{:}a = q \qquad \text{if and only if } p \not\supset a. \tag{3.2}$$

Now let $\mathfrak{S}$ be a principal element lattice or P-lattice. It is shown then in W-D [1] that

$$\text{if } a \supset b, \text{ then } b = a(b{:}a). \tag{3.3}$$

It is also shown in W-D [1] that every P-lattice is a Noether lattice; we state this result as a lemma:

LEMMA 3.4. *In a P-lattice, every irreducible element is primary.*

It will be recalled that the irreducibility here is with respect to the cross-cut operation.

4. THEOREM 4.1. *An element of a P-lattice is primary if and only if it is a power of a prime.*

Proof. Let q be primary, p its corresponding prime. Let k be the least power of p for which (3.1) holds. We prove by induction that $q = p^k$.

If $k = 1$, $q = p^1 = p$. Assume the theorem proved for all primaries q' corresponding to p such that

$$p \supset q' \supset p^s \qquad (1 \leqq s < k).$$

By Lemmas 3.2 and 3.3, $q{:}p$ is primary, corresponds to p and is not equal to q. Since $q{:}p \supset p^{k-1}$, $q{:}p = p^t$ by the hypothesis of the induction. Hence $q \supset p^{t+1}$, so that $t + 1 \geqq k$, $t \geqq k - 1$. But since $q{:}p \supset p^{k-1}$, $t \leqq k - 1$. Hence $t = k - 1$, $q{:}p = p^{k-1}$. But $p \supset q$. Hence by (3.3), $q = p(q{:}p) = pp^{k-1} = p^k$.

We need an additional lemma to prove conversely that every power of a prime is primary.

LEMMA 4.1. *If p is a prime of a P-lattice and $p^{k+1} \neq p^k$, then $p^{k+1}{:}p^k = p$.*

For we have by hypothesis, (3.3), and the trivial inclusions $p^k \supset p^{k+1} \supset p^k p$,
(i) $p^{k+1} \neq p^k$;
(ii) $p^{k+1} = p^k(p^{k+1}{:}p^k)$;
(iii) $p^{k+1}{:}p^k \supset p$.

By (iii) and (3.3), $p = (p^{k+1}{:}p^k)(p{:}(p^{k+1}{:}p^k))$. If $p^{k+1}{:}p^k \neq p$, $p \not\supset p^{k+1}{:}p^k$, so that by Lemma 3.3, $p = p{:}(p^{k+1}{:}p^k)$. Hence

$$p = (p^{k+1}{:}p^k)p. \tag{4.1}$$

Multiply both sides of (4.1) by p^{k-1} and apply (ii). Then $p^k = (p^{k+1}{:}p^k)p^k = p^{k+1}$ and (i) is contradicted. Hence $p^{k+1}{:}p^k$ must equal p.

We now prove the second part of the theorem by induction. Let p be a prime. Then $p^1 = p$ is primary. Assume that p^k is primary, $k \geqq 1$. Then p^{k+1} is primary. For assume the contrary. Then we have (i).

Furthermore, there exist lattice elements a and b such that
(iv) $p^{k+1} \supset ab$, $p^{k+1} \not\supset a$, $p^{k+1} \not\supset b^s$ $(s = 1, 2, 3, \cdots)$.
Hence $p \supset ab$, $p \not\supset b$, so that $p^k \supset ab$. Then $p^k \supset a$. For otherwise since p^k is primary and p is prime, $p^k \supset b^s$ for some s, so that $p \supset b$. If $p^k \supset a$, then $a = p^k q$ where $p \not\supset q$ by (iv). Hence $p^{k+1}{:}a = p^{k+1}{:}(p^k q) = (p^{k+1}{:}p^k){:}q = p{:}q = p$ by Lemmas 4.1 and 3.3. But since $p^{k+1} \supset ab$, $p^{k+1}{:}a \supset b$, or $p \supset b$, a contradiction. Hence p^{k+1} must be primary, and the proof is complete.

The following theorem may also be readily proved by induction:

THEOREM 4.2. *In a P-lattice, every primary element is cross-cut irreducible.*

An element s of a residuated lattice is said to be "multiplicatively irreducible" if $s = xy$ implies $s = x$ or $s = y$.

THEOREM 4.3. *In a P-lattice, an element is multiplicatively irreducible if and only if it is a prime.*

Proof. It is evident that every prime is multiplicatively irreducible. To prove conversely that every multiplicatively irreducible element is a prime, we shall need the following lemma.

LEMMA 4.2. *If for every element x of a residuated lattice $b:x = b$ or i, then b is a prime.*

For assume that $b \supset ac$ and $b \not\supset a$. Then $b:a \neq i$, so that $b:a = b$. But then $b:c = (b:a):c = b:ac = i$, so that $b \supset c$. Hence b is a prime.

Now let p be any multiplicatively irreducible element of $\mathfrak{S}$. Then if p is simple,[3] p is a prime. For since $p:x \supset p$, either $p:x = p$ or $p:x = i$ for every x of $\mathfrak{S}$, so p is a prime by Lemma 4.2.

Assume next that p is multiplicatively irreducible but not simple. Let a be any element of $\mathfrak{S}$ such that $p:a \neq i$. Then $p:a = p:d$, where $d = a \cup p \supset p$, $d \neq p$. But then by (3.3), $p = (p:d)d = (p:a)d$. Hence since p is multiplicatively irreducible, $p:a = p$. Thus p is a prime by Lemma 4.2.

COROLLARY. *If p is a prime of a P-lattice and q any proper divisor of p, then*

$$qp = p. \tag{4.4}$$

Proof. Since $q \supset p$, $q \neq p$, $p = (p:q)q$ by (3.3). But $p:q = p$.

5. We shall now give the decomposition theorems in a P-lattice. If the ascending chain condition holds in a lattice $\mathfrak{S}$, for every element we have a cross-cut representation

$$a = s_1 \cap s_2 \cap \cdots \cap s_k , \tag{5.1}$$

where each s is irreducible, and no s divides the cross-cut of the remaining s. If $\mathfrak{S}$ is distributive, it is well known that this decomposition is unique.

In a Noether lattice, every irreducible s is primary, and we obtain from (5.1) a decomposition into maximum primary components by grouping together all primaries belonging to the same prime:

$$a = q_1 \cap q_2 \cap \cdots \cap q_l. \tag{5.2}$$

Here q_i corresponds to the prime p_i, and $p_1, p_2, \cdots, p_l$ are distinct, although they may divide one another. In any other such decomposition of a, the number of primary components is the same, and the primes are the same. We shall refer to (5.2) as a "Noether decomposition" of a. The other cross-cut decompositions of a into isolated primary components and into coprime components readily follow from the Noether decomposition as in the instance of ideal theory.

The following theorem follows immediately from the identity of the primaries, irreducibles and powers of primes in a P-lattice, and the fact that every P-lattice is distributive (W-D [1], Theorem 13.2).

THEOREM 5.1. *In a P-lattice, every element a has a unique reduced decomposition into a cross-cut of powers of distinct primes:*

$$a = p_1^{\alpha_1} \cap p_2^{\alpha_2} \cap \cdots \cap p_l^{\alpha_l}. \tag{5.3}$$

[3] A lattice element is said to be "simple" if its only divisors are itself and the identity.

A multiplicative decomposition of a,

$$a = b_1 b_2 \cdots b_s,$$

is said to be reduced if for each b_i, $b_1 b_2 \cdots b_{i-1} b_{i+1} \cdots b_s \neq a$. This definition implies that no $b = a$, so that a multiplicatively irreducible element will admit of no such decomposition.

THEOREM 5.2. *In a P-lattice, every element has a unique reduced multiplicative decomposition into a product of powers of distinct primes:*

$$a = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_l^{\alpha_l}. \tag{5.4}$$

Proof. We shall prove by induction that

$$p_1^{\alpha_1} \cap p_2^{\alpha_2} \cap \cdots \cap p_l^{\alpha_l} = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_l^{\alpha_l} \tag{5.5}$$

if $p_1, p_2, \cdots, p_l$ are distinct primes of $\mathfrak{S}$. The theorem will then follow from Theorem 5.1.

(5.5) is true when $l = 1$. Assume that it is true for any $l - 1$ distinct primes of $\mathfrak{S}$. Select from $p_1, p_2, \cdots, p_l$ a prime not divisible by any other prime; let it be p_1. Then if $a = p_1^{\alpha_1} \cap p_2^{\alpha_2} \cap \cdots \cap p_l^{\alpha_l}$, $p_1^{\alpha_1} \supset a$. Hence by (3.3), $a = p_1^{\alpha_1}(a:p_1^{\alpha_1})$.

Now $a:p_1^{\alpha_1} = (p_1^{\alpha_1} \cap p_2^{\alpha_2} \cap \cdots \cap p_l^{\alpha_l}):p_1^{\alpha_1} = p_1^{\alpha_1}:p_1^{\alpha_1} \cap p_2^{\alpha_2}:p_1^{\alpha_1} \cap \cdots \cap p_l^{\alpha_l}:p_1^{\alpha_1}$. But $p_1^{\alpha_1}:p_1^{\alpha_1} = i$, and by Lemma 3.3, $p_j^{\alpha_j}:p_1^{\alpha_1} = p_j^{\alpha_j}$ $(j \neq 1)$. Hence $a:p_1^{\alpha_1} = p_2^{\alpha_2} \cap \cdots \cap p_l^{\alpha_l} = p_2^{\alpha_2} \cdots p_l^{\alpha_l}$ by the hypothesis of the induction. Thus $a = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_l^{\alpha_l}$, and the proof is complete.

THEOREM 5.3. *There exist distributive Noether lattices in which the Noether decomposition of an element into a cross-cut of primary components is not unique.*

Proof. Consider the lattice $\mathfrak{L}$ (Figure 1) with the following multiplication table

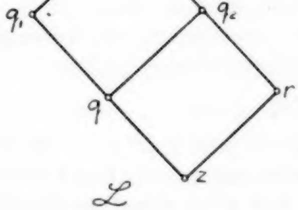

FIG. 1

$x \backslash y$	i	p	q_1	q_2	q	r	z
i	i	p	q_1	q_2	q	r	z
p	p	q	q	q	q	z	z
q_1	q_1	q	q	q	q	z	z
q_2	q_2	q	q	q	q	z	z
q	q	q	q	q	q	z	z
r	r	z	z	z	z	z	z
z	z	z	z	z	z	z	z

It is easily verified that the postulates for a multiplication given in W-D [1], pp. 336–337 are all satisfied, so that $\mathfrak{L}$ is residuated. An inspection of the multiplication table shows that p and r are primes in $\mathfrak{L}$ while q_1, q_2, q are primaries corresponding to p. Since p, q_1, q_2, r are the only irreducible elements, $\mathfrak{L}$ is a Noether lattice. It is evidently distributive. However, the element z has two distinct decompositions into primary components, $z = q_1 \cap r$, $z = q \cap r$.

If the decomposition (5.1) into cross-cut irreducible elements is unique, severe conditions are imposed on the lattice structure; in fact the lattice must be a Birkhoff lattice of a peculiar type (Dilworth [1]). On the other hand, we may have a unique Noether decomposition (5.2) for every lattice element with hardly any restriction on the lattice. For it is easily seen that a sufficient condition for a unique Noether decomposition (5.2) of a given lattice element a of an N-lattice is that none of the primes belonging to a divide one another. Hence in a Noether lattice in which every prime is simple, the Noether decomposition is always unique. But such a lattice is the direct cross-cut of Noether lattices containing exactly one simple element. Now it is shown in W-D [1], Theorem 8.1 that a lattice with only one simple element may always be residuated. If the lattice is finite with null element z, the corresponding multiplication is easily shown to be $ab = a$ if $b = i$; $ab = b$ if $a = i$; $ab = z$ otherwise. Hence every element save i is primary, and the lattice is a Noether lattice. Since the lattice is otherwise arbitrary, we may construct by direct cross-cuts finite Noether lattices with unique decomposition of almost any desired complexity of lattice structure; in particular, the lattice need not be distributive, modular, nor a Birkhoff lattice.

6. THEOREM 6.1. *Necessary and sufficient conditions that an N-lattice be a P-lattice are*

(i) *every primary is a power of a prime;*

(ii) *if p is a prime and q a proper divisor of p, then $qp = p$.*

Proof. The necessity of these conditions is established in §4. Let $\mathfrak{S}$ be an N-lattice in which (i) and (ii) hold, and let a and b be any two elements of it such that $a \supset b$. It suffices to exhibit an element c such that $ac = b$.

Noether decompositions of a and b exist of the forms

$$a = p_1^{\alpha_1} \cap p_2^{\alpha_2} \cap \cdots \cap p_l^{\alpha_l}, \qquad b = s_1^{\beta_1} \cap s_2^{\beta_2} \cap \cdots \cap s_k^{\beta_k},$$

where $p_1, \cdots, p_l$ are distinct primes and $s_1, \cdots, s_k$, distinct primes.

As in the proof of Theorem 5.2, we may show by induction that these decompositions imply the existence of multiplicative decompositions of the form

$$a = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_l^{\alpha_l}, \qquad b = s_1^{\beta_1} s_2^{\beta_2} \cdots s_k^{\beta_k}.$$

Since $a \supset b$, each p divides some s. But if $p \supset s$, $p \neq s$, $ps = s$. Hence $p^{\alpha} s = s$, $p^{\alpha} s^{\beta} = s^{\beta}$. If $p \supset s$, $p = s$, then if s is the only prime which belongs to b which p divides, $p^{\alpha} \supset s^{\beta}$ or $p^{\alpha} \supset p^{\beta}$. If $\alpha \geqq \beta$, $p^{\alpha} = p^{\beta}$ so that we may replace p^{α} by p^{β} in a. If $\alpha < \beta$, $p^{\beta-\alpha} p^{\alpha} = p^{\beta} = s^{\beta}$. If p divides some other

prime in b, say s', by the first case $p^\alpha s'^\beta = s'^\beta$. The element c is thus constructed by selecting for each p the power s^β of which it is a divisor (save in the case $p = s$ when we select the power $s^{\beta-\alpha}$ or $s^0 = i$ according as $\beta > \alpha$ or $\beta \leqq \alpha$), and then taking the product of these powers.

III. Semi-arithmetical lattices

7. Let $\mathfrak{S}$ now denote a residuated lattice with ascending chain condition in which any one of the three following conditions hold:

$$a:b \cup b:a = i; \qquad a:(b \cap c) = a:b \cup a:c; \qquad (b \cup c):a = b:a \cup c:a. \tag{7.1}$$

It is proved in W-D [1] that these three conditions are equivalent to one another and imply that $\mathfrak{S}$ is distributive. It was also proved in W [2] that a sufficient condition for (7.1) is

$$a:(a:b) = a \cup b. \tag{7.2}$$

THEOREM 7.1. *Let $\mathfrak{S}$ be a residuated lattice with ascending chain condition in which conditions* (7.1) *hold. Then any two primaries of $\mathfrak{S}$ are either coprime or else one contains the other.*

Proof. Let p and q be any two primes of $\mathfrak{S}$. By Lemmas 3.2 and 3.3, $p:q = p$ or i, $q:p = q$ or i. By (7.1), $p:q \cup q:p = i$. Hence any two primes of $\mathfrak{S}$ are either coprime or else one divides the other.

Now let a and b be any two primaries of $\mathfrak{S}$ such that neither $a \supset b$ nor $b \supset a$. Then a and b are coprime. For otherwise, there exists a prime r such that $r \supset a \cup b$, so that $r \supset a$, $r \supset b$. Hence if p and q are the primes of Lemma 3.1 which correspond to a and b respectively, $r \supset p$, $r \supset q$. Hence $p \supset q$ or $q \supset p$. But by Lemma 3.2, $p \supset a:b$, $q \supset b:a$. Hence if $p \supset q$, $p \supset a:b$, $p \supset b:a$, and this contradicts (7.1). Similarly $q \supset p$ leads to a contradiction of (7.1), and the proof is complete.

The following simple example shows that the conditions (7.1) do not imply that $\mathfrak{S}$ is a Noether lattice. Consider a chain of four elements i, a, b, z with $i > a > b > z$. Define a multiplication xy by $xy = yx$; $xy = x$ if $y = i$; $xz = z$, $aa = a$, $ab = bb = z$. This lattice is residuated. (The values of $x:y$ are given in W-D [1], p. 339, example III.) Furthermore since either $x \supset y$ or $y \supset x$, $x:y \cup y:x = i$, so that (7.1) holds. Every element is cross-cut irreducible. But $z \supset ab$, $z \not\supset b$, $z \not\supset a^n$ for any n, so that z is non-primary.

DEFINITION. *A Noether lattice in which conditions* (7.1) *hold is said to be semi-arithmetical.*

We refer to such a lattice as an *SA*-lattice.

THEOREM 7.2. *A Noether lattice is semi-arithmetical if and only if any two primaries of the lattice are either coprime, or else one divides the other.*

Proof. The necessity of the condition is proved in Theorem 7.1. To show its sufficiency, let a and b be any two elements of a Noether lattice $\mathfrak{S}$. Then

a and b have Noether decompositions into cross-cuts of primaries. But since of any two primaries with a common factor one divides the other, the primaries in any reduced representation are coprime in pairs, so that the cross-cut representation is also a product representation. We may accordingly write

$$a = q_1 \cap \cdots \cap q_k \cap s_1 \cap \cdots \cap s_l = q_1 \cdots q_k s_1 \cdots s_l ,$$
$$b = q_1' \cap \cdots \cap q_k' \cap t_1 \cap \cdots \cap t_m = q_1' \cdots q_k' t_1 \cdots t_m ,$$

where the q, s, q', t are primaries and

$$q \cup s = q \cup t = s \cup q' = s \cup t = i,$$
$$q_i \cup q_j = i, \qquad q_i' \cup q_j' = i \qquad\qquad (i, j = 1, \cdots, k; i \neq j),$$
$$s_u \cup s_v = i, \qquad t_p \cup t_w = i \; (u, v = 1, \cdots, l; u \neq v; p, w = 1, \cdots, m; p \neq w),$$

while either $q_i \supset q_i'$ or $q_i' \supset q_i$.

Then by elementary properties of the residual

$$a:b = q_1:b \cap \cdots \cap s_l:b, \qquad q_1:b = q_1:(q_1'\bar{q}) = (q_1:\bar{q}):q_1' ,$$

where $\bar{q} = q_2' \cdots q_k' t_1 \cdots t_m$. Since q_1 is primary and relatively prime to each of $q_2', \cdots, q_k', t_1, \cdots, t_m$, it is relatively prime to their product $\bar{q}$. Hence by Lemma 3.2, $q_1:\bar{q} = q_1$ so that $q_1:b = q_1:q_1'$. In like manner, $q_2:b = q_2:q_2', \cdots, q_k:b = q_k:q_k'$, while $s_1:b = s_1, \cdots, s_l:b = s_l$. Hence

$$a:b = q_1:q_1' \cap \cdots \cap q_k:q_k' \cap s_1 \cap \cdots \cap s_l .$$

In like manner

$$b:a = q_1':q_1 \cap \cdots \cap q_k':q_k \cap t_1 \cap \cdots \cap t_m .$$

Since either $q':q = i$ or $q:q' = i$, the primaries appearing in these decompositions of $a:b$ and $b:a$ are coprime in pairs. Hence $a:b \cup b:a = i$ and $\mathfrak{S}$ is semi-arithmetical.

THEOREM 7.3. *In a semi-arithmetical lattice, every primary element is irreducible.*

Proof. Let q be primary, and let p be its corresponding prime. If q is reducible, we have a reduced decomposition

$$q = q_1 \cap q_2 \cap \cdots \cap q_k , \qquad k \geqq 2,$$

where the q_i are irreducible and no q_i divides the cross-cut of the remainder. Since the q_i are primary, they are accordingly relatively prime in pairs. Hence if $\bar{q}_1 = q_2 \cap \cdots \cap q_k$, $q \supset q_1\bar{q}_1$, $q \not\supset \bar{q}_1$. Hence $q \supset q_1^s$ for some integer s. Therefore, $p \supset q_1$. Similarly $p \supset q_2$, and this contradicts the coprimeness of q_1 and q_2.

THEOREM 7.4. *In a semi-arithmetical lattice $\mathfrak{S}$, every power of a prime is cross-cut irreducible and primary.*

Proof. Let p be a prime of $\mathfrak{S}$ and $q = p^k$ any power of p. If q were reducible, we would have with $l \geqq 2$

$$q = q_1 \cap q_2 \cap \cdots \cap q_l = q_1 q_2 \cdots q_l,$$

where the q_i are primary and coprime in pairs, and no q_i equals q. Clearly p contains exactly one q_i; let it be q_1, and let $\bar{q} = q_2 \cdots q_l$. Then $p \cup \bar{q} = i$. Hence $q \cup \bar{q} = i$. Hence $q:\bar{q} = q$, so that $q:q_1 = (q:\bar{q}):q_1 = q:\bar{q}q_1 = q:q = i$, a contradiction.

These results give us the following decomposition theorem.

THEOREM 7.4. *In a semi-arithmetical lattice, every element may be uniquely expressed either as a cross-cut or a product of irreducible elements which are coprime in pairs.*

8. The structure of a semi-arithmetical lattice is largely determined by the structure of the semi-ordered set of its prime elements. For consider a chain of elements of

$$\cdots, q, \cdots, q', \cdots, q'', \cdots, \tag{8.1}$$

where each q is primary and contains all the q which follow it. Since the ascending chain condition holds in $\mathfrak{S}$, any such chain is well-ordered. For convenience, we shall assume that the chain begins with the identity element of $\mathfrak{S}$, and terminates with the null element, if the latter exists. We shall furthermore assume that all the elements of the chain are distinct, and that the chain is complete in the following sense: If q' is the immediate successor of q in the chain and $q \supset b \supset q'$ in the lattice, then either $q = b$ or $q' = b$ or b is non-primary. We call such a chain a "complete chain of primaries" of $\mathfrak{S}$. Every element of such a chain is cross-cut irreducible by Theorem 7.3.

The following properties of complete chains of primaries follow easily from the definition:

(i) The second element of the chain is a simple element of $\mathfrak{S}$ and hence a prime.

(ii) If q lies in the chain and b is any primary contained in q, b lies in the chain.

(iii) Every such chain is a residuated sublattice of $\mathfrak{S}$ and hence a semi-arithmetical lattice.

(iv) Two such chains either have no element in common save the identity, or else coincide completely from a certain point on.

(v) An element of a chain is a prime if and only if it has at least one prime divisor distinct from the prime divisors of all its predecessors in the chain.

(vi) Two primaries of the chain with the same set of prime divisors correspond to the same prime in the sense of Lemma 3.1, and this prime belongs to the chain.

(vii) A non-simple element is a prime if and only if it belongs to more than one complete chain of primaries.

Consequently the first element other than the identity common to two complete chains of primaries is always a prime.

(viii) If an element of the chain is a prime, all the primaries corresponding to it belong to the chain, and form with the identity element a semi-arithmetical sublattice of $\mathfrak{S}$.

It follows from the decomposition theorems that the residuation in a semi-arithmetical lattice is completely determined if it is determined for any two primary elements of the lattice. Now if q and q' are primary, either they are coprime or else they belong to a complete chain of primaries. If they are coprime, $q:q' = q$, $q':q = q'$ by Lemma 3.3. If $q' \supset q$, $q':q = i$. On the other hand, $q:q' = q$ if $q' \supset q$ and q, q' correspond to different primes in the chain. Hence we see that the residuation over $\mathfrak{S}$ is completely determined if it is specified in those sublattices of $\mathfrak{S}$ consisting of all the primaries corresponding to a given prime and the identity. The distribution of these sublattices in $\mathfrak{S}$ is given by the semi-ordered set of all prime elements of $\mathfrak{S}$, in that each sublattice corresponds to a covering of prime elements in the semi-ordered set.

9. We shall conclude by showing that in an Archimedean lattice, condition (7.2) (whose combinatorial consequences we developed in the earlier paper W [2]) determines the residuation of the lattice uniquely.

For consider the sublattice consisting of i and all the primaries belonging to a given prime p. Since $\mathfrak{S}$ is Archimedean, the sublattice may be arranged in a sequence

$$q_0, q_1, q_2, \cdots, q_n, q_{n+1}, \cdots,$$

where $q_0 = i$, $q_1 = p$ and q_n covers q_{n+1} if $n \geqq 1$. The sequence terminates with q_N if and only if there is a finite number N of primaries belonging to p.

Consider any primary q_k of the sequence. Evidently $q_k:q_l = q_0$ if $l \geqq k$. The $k + 1$ residuals $q_k:q_0, q_k:q_1, \cdots, q_k:q_k$ are all distinct. For if $l, m \leqq k$, then by (7.2) $q_l = q_l \cup q_k = q_k:(q_k:q_l)$, $q_m = q_m \cup q_k = q_k:(q_k:q_m)$. Hence if $q_k:q_l = q_k:q_m$, $q_l = q_m$ or $l = m$. But each residual $q_k:q_l$ contains q_k and if $l < k$, it is primary and belongs to p by Lemma 3.2. Also if $q_l \supset q_m$, $q_k:q_m \supset q_k:q_l$. Hence if $l \leqq k$, $q_k:q_l = q_{k-l}$. Thus we have in general

$$q_k:q_n = \begin{cases} q_{k-n} & \text{if } k - n \geqq 0, \\ q_0 & \text{if } k - n \leqq 0, \end{cases}$$

so that the residuation in the sublattice is determined. By W-D [1], Theorem 5.1, the associated multiplication is found to be

$$q_k q_l = q_{k+l},$$

if the number of primaries is infinite, and if the number is finite and equal to N, $q_k q_l = q_{k+l}$ or q_N according as $k + l \leqq N$ or $k + l \geqq N$. Hence if we write $p^0 = i$, it follows that $q_k = p^k$, so that every primary is a power of a prime.

The residuation in the remainder of the lattice is then determined by the results of §8.

References

R. P. Dilworth.

1. *Lattices with unique irreducible decomposition.* To appear shortly in the Annals of Mathematics.

M. Ward.

1. *Structure residuation,* Annals of Mathematics, vol. 39(1938), pp. 558-568.
2. *Residuation in structures over which a multiplication is defined,* this Journal, vol. 4(1937), pp. 627-636.

M. Ward and R. P. Dilworth.

1. *Residuated lattices,* Transactions of the American Mathematical Society, vol. 45(1939), pp. 335-354.
2. *The lattice theory of ova,* Annals of Mathematics, vol. 40(1939), pp. 600-608.

California Institute of Technology.

FUNDAMENTAL SYSTEMS OF SOLUTIONS FOR LINEAR DIFFERENCE EQUATIONS

By M. G. Moore

Consider the differential equation with constant coefficients

$$\sum_{\mu=0}^{n} c_\mu f^{(\mu)}(x) = 0.$$

If the function exp (tx) is to be a solution, its substitution into the equation shows that t must then be a zero of the polynomial

$$\sum_{\mu=0}^{n} c_\mu t^\mu.$$

Further, with each of the zeros is associated an exponential function which satisfies the differential equation, so that, in general, there exist n distinct exponential solutions. Also, all possible solutions of the equation may, in general, be expressed in terms of these exponentials, and simply as linear combinations.

Consider now the generalized difference equation

$$\sum_{\mu=1}^{N} c_\mu f(x + \alpha_\mu) = 0, \tag{1}$$

with constant coefficients and with complex spans α_μ. Again seeking solutions of the form exp (tx), we substitute into (1) and find that t must be a zero of the function

$$h(t) = \sum_{\mu=1}^{N} c_\mu \exp(\alpha_\mu t).$$

On the other hand, if t is a zero of $h(t)$, exp (tx) is a solution of (1).

It is natural to inquire now into the possibility that the analogy between these two functional equations can be extended to a consideration of the "general" solution of (1).

That the analogy cannot be perfect is immediately evident when we consider the particular equation

$$f(x + 1) - f(x) = 0$$

and notice that the number of its exponential solutions (exp $(2m\pi ix)$, m integral) is infinite. In this case, then, not only do we find solutions by taking linear combinations of these exponentials, but it is possible to arrive at still others by taking the limits of converging sequences of such linear combinations. The problem of how extensive a class of solutions is thus obtained is exactly the problem which is attacked by the best-known part of the theory of Fourier series. If we use the well-known set of Fourier orthogonality relations as the

Received January 30, 1940; presented to the American Mathematical Society, April 15, 1939.

basis of its method for obtaining coefficients, the Fourier theory deals with the problem so effectively as to make it clear that these properties of orthogonality are among the most powerful tools of mathematics.

If we are to deal with the solutions of (1) with a degree of success comparable to that already attained for periodic functions, it seems likely that we shall need to use closely analogous methods; that is, that we must employ relations of orthogonality, or of biorthogonality, satisfied by the exponential solutions of (1). It would then be possible to assign coefficients to very general types of solutions of the equation. There should be some expectation of being able to utilize such coefficients in the study of the possibility of approximation.

Such properties of biorthogonality have actually been obtained by the author,[1] and will be applied here to the treatment of the question of the existence of a "fundamental system" of exponential solutions[2] when the function $h(t)$ has only simple zeros, while the corresponding situations which involve multiple zeros will be treated with equal success.

Although the spans α_μ are not restricted to be real, we may take them so, and the function to be approximated will still be defined on curves (or in regions) in the complex plane, not necessarily coinciding with the real axis.[3]

Let c_μ and α_μ ($\mu = 1, 2, \cdots, N$) be arbitrary complex constants except that $c_\mu \neq 0$, $\alpha_\mu \neq \alpha_\nu$ for $\mu \neq \nu$, and $N \geqq 2$. Let $w(m)$ be the order of the zero t_m of $h(t)$. We have methods then for obtaining formal expansions of functions in terms of the solutions $x^q \exp(t_m x)$ ($q = 0, 1, 2, \cdots, w(m) - 1$) of (1). Sufficient conditions for the convergence of these formal expansions (conveniently called F-series) are given in (M), where the partial sum of the F-series of an "arbitrary" function $f(x)$ is given in the form (with the point a at our disposal)

$$\frac{1}{2\pi i}\sum_{\mu=1}^{N} c_\mu \int_a^{\alpha_\mu} f(x_1) \int_{C_s} e^{(\alpha_\mu + x - x_1)t}\{h(t)\}^{-1}\, dt\, dx_1.$$

Here C_s ($s = 1, 2, \cdots$) is a contour which, except for a bounded portion of its length, is a circle of radius s, and on which

$$| e^{-\alpha_\mu t} h(t) | > m_1 \qquad (\mu = 1, 2, \cdots, N),$$

m_1 being a positive quantity independent of s.[4]

[1] M. G. Moore, *On expansions in series of exponential functions*, American Journal of Mathematics, vol. 62(1940), pp. 83–90. We refer to this paper as (M).

[2] See G. Valiron, *Sur les solutions des équations différentielles linéaires d'ordre infini et à coefficients constants*, Annales Scientifiques de l'Ecole Normale Supérieure, (3), vol. 46(1929), pp. 25–53, for a study of the fundamental solutions of differential equations of infinite order. The solutions which he considers are analytic in certain circles.

[3] For the case where they do coincide with the real axis, see T. Kitagawa, *On the theory of linear translatable functional equations and Cauchy's series*, Japanese Journal of Mathematics, vol. 13(1936), pp. 233–332. Kitagawa also refers to various other authors.

[4] For the real problem, the F-series is a special case of Kitagawa's Cauchy series. See Kitagawa, loc. cit. For other properties of $h(t)$, see R. E. Langer, *On the zeros of exponential sums and integrals*, Bull. Am. Math. Soc., vol. 37(1931), pp. 213–239. See also R. D. Carmichael, *Systems of linear difference equations and expansions in series of exponential functions*, Trans. Am. Math. Soc., vol. 35(1933), pp. 1–28.

We may suppose that $h(0) \neq 0$, for otherwise, if γ is any constant for which $h(-\gamma) \neq 0$, we may replace $h(t)$ by $H(t) \equiv h(t-\gamma)$, and $f(x)$ by $F(x) \equiv f(x) \exp \gamma x$, and our new problem of approximation is equivalent to the old one.

Let the complex numbers η_k $(k = 1, 2, \cdots, M)$ be so defined that $\eta_k = \eta e^{i\delta_k}$ (η positive, δ_k real), let R_k be an infinite ray projecting from the origin, but not including the origin, and let R_k and η_k be so defined that $R(\eta_k t)$ (R denoting the real part) is positive for t on R_k and R_{k+1}, where $R_{M+1} = R_1$. Also, let R_k pass through no zero of $h(t)$, and let it be different from all rays on which $R(\alpha_\mu t) = R(\alpha_\nu t)$ $(\mu \neq \nu)$. Let C_{ks} be the contour made up of portions of R_k, C_s, and R_{k+1}. Then $| e^{-\alpha_\mu t} h(t) |$ $(\mu = 1, 2, \cdots, N)$ is bounded away from zero on C_{ks}, the bound being independent of k and s.

Let $\alpha^{(k)}$ be the value α_μ $(\mu = 1, 2, \cdots, N)$ for which $R(\alpha^{(k)} t) > R(\alpha_\nu t)$ $(\alpha^{(k)} \neq \alpha_\nu)$ for t on R_k.

Let χ be arbitrary in each term of the sum $\Sigma_1 \equiv \Sigma(\alpha_\chi - \alpha^{(k)})$, the number of terms also being arbitrary. Let x_1 lie on a rectifiable arc from x to α_μ on which $f(x_1)$ is defined and to which we shall refer as the arc (x, α_μ). Let all possible points $\alpha_\mu - \alpha^{(k)} + x - x_1 + \Sigma_1$ lie in an angle with vertex at the origin and lying interior to the reflex angle formed by the rays from the origin through the points η_k and η_{k-1}, and let $\alpha_\mu - \alpha^{(k)} + x - x_1 + \Sigma_1$ also be bounded away from the origin except when $\alpha_\mu - \alpha^{(k)} + \Sigma_1 = 0$.

THEOREM 1. *Let the above hypotheses hold and let $f(x_1)$ have a continuous second derivative on closed (x, α_μ) $(\mu = 1, 2, \cdots, N)$. Then,*

$$\lim_{\eta \to 0} \lim_{s \to \infty} \frac{1}{2\pi i} \sum_{k=1}^{M} \sum_{\mu=1}^{N} \int_a^{\alpha_\mu} f(x_1) \int_{C_{ks}} \frac{c_\mu e^{(\alpha_\mu + x - x_1)t}}{h(t)\Gamma(1+\eta_k t)} \, dt \, dx_1 = f(x). \tag{2}$$

Proof. Any portion of R_k $(k = 1, 2, \cdots, M)$ of bounded length contributes zero to the value of (2) as $\eta \to 0$ because of a cancellation of integrals having opposite directions over R_k, so that we shall neglect the portion of R_k lying within a distance T (T to be independent of s and η) of the origin. We define D_{ks} to be the remaining part of C_{ks} after we have neglected this portion of R_k and the corresponding portion of R_{k+1}, and we take T_k as the point on R_k for which $| T_k | = T$.

The left side of (2) may then be written, upon repeated integration by parts, as

$$\begin{aligned} \lim_{\eta \to 0} \lim_{s \to \infty} \frac{1}{2\pi i} \sum_{k=1}^{M} \int_{D_{ks}} \Bigg[& \frac{f(x)}{t} + \frac{f'(x)}{t^2} - (h(t))^{-1} \\ & \cdot \sum_{\mu=1}^{N} c_\mu \left(\frac{f(\alpha_\mu) e^{xt}}{t} - \frac{f'(\alpha_\mu) e^{xt}}{t^2} - \int_x^{\alpha_\mu} \frac{f''(x_1) e^{(\alpha_\mu + x - x_1)t}}{t^2} \, dx_1 \right) \Bigg] \frac{dt}{\Gamma(1+\eta_k t)}. \end{aligned} \tag{3}$$

The first term in the bracket, by Cauchy's integral formula, contributes $f(x)$ to the value of (3).

The arc of C_s on D_{ks} contributes zero for the values of each of the other terms because of the presence of $\Gamma(1 + \eta_k t)$ in the denominator and the boundedness of $| e^{\alpha_\mu t}/h(t) |$ for all μ and for all t on C_s for all s.

If we write $\{h(t)/c^{(k)}e^{\alpha^{(k)}t}\}^{-1}$ as the reciprocal of a binomial $1 + [h(t)(c^{(k)}e^{\alpha^{(k)}t})^{-1} - 1]$ and expand by the binomial theorem for sufficiently large t on R_k, we find that in the last term of (3) we may obtain an expansion of $e^{(\alpha_\mu+x-x_1)t}/h(t)$, a representative term of which (except for a constant coefficient) is $z(t) = \exp\{[\alpha_\mu - \alpha^{(k)} + x - x_1 + \Sigma_1]t\}$. Further, by taking T sufficiently large, we find that, except for a finite number of terms, the expansion not only is uniformly convergent when $|t| > T$ and t is on R_k, but every term approaches zero in an exponential manner as $T \to \infty$. We need consider, then, only a finite number of terms of the form $z(t)$. By the hypotheses of the theorem, there exists a T_k' ($|T_k'| = T$) for which $|z(T_k')| < 1$, $R(\eta_k T_k') > 0$, and $R(\eta_{k-1}T_k') > 0$ (the hypotheses having been chosen to bring this about). The integral which remains, then, for each of the finite number of terms which do not approach zero for large T, can be replaced by an integral from T_k to T_k', and from T_k' to C_s on a ray from the origin. The first part of the path of integration, being bounded in length, contributes zero to the value of the integral, while on the second, the value of the integral becomes zero with $1/T$ because of the presence of t^2 in the denominator and the boundedness of other factors. The other terms of (3) are treated similarly, the $z(t)$ (in those cases) approaching zero exponentially for large t on the ray from the origin through T_k'.

THEOREM 2. *There exist values η_k and a rectifiable arc A passing through the points $\alpha_1, \alpha_2, \cdots, \alpha_N$ such that, if $f(x_1)$ is defined for x_1 on A, and if x is on A but is a point α_μ for no value of μ, the hypotheses of Theorem 1 are satisfied.*

Proof. Take any line L through the origin not parallel to any one of the finite set of lines joining the points α_μ and α_ν ($\mu \neq \nu$).

Let 5ϵ be a positive quantity smaller than each of the acute angles which L makes with these lines. Then if δ is a real number which is defined so that the point $e^{i\delta}$ lies on L, take $\delta_1 = \delta - 3\epsilon$, $\delta_2 = \delta + \epsilon$, $\delta_3 = \delta + \pi - 3\epsilon$, $\delta_4 = \delta + \pi + \epsilon$, while the arguments of points of R_1, R_2, R_3, R_4 respectively are to be $-\delta + \frac{1}{2}\pi + 2\epsilon$, $-\delta + \frac{1}{2}\pi - 2\epsilon$, $-\delta - \frac{1}{2}\pi + 2\epsilon$, $-\delta - \frac{1}{2}\pi - 2\epsilon$.

Now take A to be an arc joining the points α_μ ($\mu = 1, 2, \cdots, N$) and with a continually turning tangent which makes an acute (or right) angle $\geqq 4\epsilon$ with L at all points of A. We suppose that the notation is so chosen that $\alpha_1 = \alpha^{(1)}$ and that α_1 is one end-point of the arc, α_N is the other end-point, and $\alpha_2, \cdots, \alpha_{N-1}$ are distributed in order between these two points. The reader may now complete the proof of the theorem without difficulty.

The nature of the hypotheses on A clearly enables us to state

THEOREM 3. *There exists a rectifiable arc K containing the points $(\alpha_1, \alpha_2, \cdots, \alpha_N)$ such that, if $f(x_1)$ is defined for x_1 on K, α_1 lies on an interval B of K on which $\sum_{\mu=1}^{N} c_\mu f(x_1 - \alpha_1 + \alpha_\mu)$ is defined, and for which $(\alpha_1 + c, \alpha_N + c)$ is an arc A for all values of c for which $\alpha_1 + c$ is on B.*

THEOREM 4. *Let the hypotheses of Theorem 1 hold and let*

$$\sum_{\mu=1}^{N} c_\mu f(x_1 + \alpha_\mu - \alpha_1) = 0 \tag{4}$$

for x_1 on B. Then the limit expressed in (2) *exists uniformly for $x - \alpha_\mu + \alpha_1$ ($\mu = 1, 2, \cdots, N$) on any closed interval of open B.*

Proof. The expression (2) is already known to be independent of a, for a on K. Now, the substitution of $\alpha_\mu + c$ for α_μ (c a constant such that $\alpha_1 + c$ lies in open B) changes the value of (2) by an amount which is conveniently written as the limit of a linear combination of terms of the form

$$\sum_{\mu=1}^{N} \int_{\alpha_\mu}^{\alpha_\mu+c} c_\mu f(x_1) \int_{\Gamma_m} \frac{e^{(\alpha_\mu+x-x_1)t}}{h(t)\Gamma(1+\eta_k t)}\, dt\, dx_1$$
$$= \int_0^c \sum_{\mu=1}^{N} c_\mu f(x_1 + \alpha_\mu) \int_{\Gamma_m} \frac{e^{(x-x_1)t}}{h(t)\Gamma(1+\eta_k t)}\, dt\, dx_1,$$

which vanish, since $\sum_{\mu=1}^{N} c_\mu f(x_1 + \alpha_\mu) = 0$.

For x on any closed interval Q (of open B) for which the distance between no two points is as great as $|\alpha_1 - \alpha_2|$, there evidently exists a c such that when the points α_μ are replaced by the points $\alpha_\mu + c$, the interval Q lies between $\alpha_1 + c$ and $\alpha_N + c$ on K and contains no point $\alpha_\mu + c$ ($\mu = 1, 2, \cdots, N$).

The reader may now see for himself that all limiting processes hold uniformly for x on Q. Any closed interval B_1 of open B is the sum of a finite number of intervals Q, so that the convergence is uniform for x on B_1. The proof for $x - \alpha_\mu + \alpha_1$ ($\mu = 2, \cdots, N$) on B_1 is similar.

Noting now that continuous functions may be approximated uniformly by sequences of functions having continuous second derivatives, and that integrable functions and various subclasses of the class of integrable functions may be approximated in various ways by functions which are continuous, we state the following theorem, which may be given more precise meaning by the reader:

THEOREM 5. *Let $f(x)$ be defined and integrable for $x - \alpha_\mu + \alpha_1$ ($\mu = 1, 2, \cdots, N$) on B and let* (4) *hold for x on B. Then $f(x)$ may be approximated by linear combinations of the solutions $x^q e^{t_m x}$ of* (1).

For the difference equation $f(x + 1) - f(x) = 0$, it is seen that the solutions may in particular be approximated by trigonometric functions if the curve on which $f(x)$ is defined has a continually turning tangent and a bounded slope. The reader may see for himself that the slope can also have discontinuities if suitable restrictions are substituted for that of continually turning tangent.

As a particular case of the results expressed in the preceding theorems, it is worth while to mention the result:

THEOREM 6. *Let $f(x)$ be defined and analytic in N regions G_μ ($\mu = 1, 2, \cdots, N$) defined so that G_μ contains the point α_μ on its interior, and let* (4) *hold for x in G_μ.*

Then there exist subregions of G_μ (including α_μ) in which $f(x)$ may be approximated uniformly by linear combinations of the solutions $x^q e^{t_m x}$ of (1).

Referring to (M) and to the definition of the F-series, we find other results by methods similar to those already used. We state them without proof.

THEOREM 7. *Let $f(x)$ be analytic throughout an open region U containing the polygon P (the smallest closed convex polygon containing $\alpha_1, \cdots, \alpha_N$, and let* (1) *be satisfied for x in a neighborhood of the origin. Let $\bar{P}$ be the smallest closed convex polygon containing a set of points β_μ which are so defined that $\beta_\mu - \alpha_\mu + \alpha_\nu$ is a point of $\bar{P}$* $(\mu, \nu = 1, 2, \cdots, N)$. *Then if $\bar{P}$ contains P and is contained in U, the F-series converges uniformly to $f(x)$ in $\bar{P}$.*

THEOREM 8. *Let $f(x)$ be an integral function which satisfies* (1). *Then its F-series converges to $f(x)$ throughout the finite plane, the convergence being uniform in any closed region.*

INDIANA UNIVERSITY.

SUFFICIENT CONDITIONS FOR VARIOUS DEGREES OF APPROXIMATION BY POLYNOMIALS

By J. L. Walsh and W. E. Sewell

1. **Introduction.** Let E be a closed limited point set in the plane of the complex variable $z = x + iy$, and let the function $f(z)$ be defined on E. Let $\{p_n(z)\}$ be a sequence of polynomials in z of respective degrees n,[1] whose convergence to $f(z)$ is to be considered. The study of the relation between regions of convergence and geometric degree of convergence to $f(z)$ on E of the sequence $p_n(z)$ on the one hand, and regions of analyticity of $f(z)$ on the other hand, is in a relatively satisfactory state in the literature.[2] But the more delicate study of the relation between the degree of convergence both on E and elsewhere of the sequence $p_n(z)$ on the one hand, and the continuity properties of $f(z)$ *on the boundary* of the region of convergence on the other hand, has only recently been undertaken. It is the object of the present paper to contribute to this latter study.

To be more explicit, let the complement (with respect to the extended plane) K of E be connected, and regular in the sense that K possesses a Green's function $G(x, y)$ with pole at infinity.[3] Then the function $w = \varphi(z) = e^{G(x,y)+iH(x,y)}$, where $H(x, y)$ is conjugate to $G(x, y)$ in K, maps K conformally (not necessarily uniformly) onto the exterior of the unit circle γ in the w-plane so that the points at infinity in the two planes correspond to each other. We denote by C the boundary of E and also denote generically by C_ρ the locus $G(x, y) = \log \rho > 0$ or $|\varphi(z)| = \rho > 1$ in K. This notation is convenient in separating the study of approximation by polynomials into several problems.

Let the function $f(z)$ be assumed analytic merely in the interior points of E; *Problem α is the study of the relation between degree of convergence of $p_n(z)$ to $f(z)$ on E on the one hand, and continuity properties of $f(z)$ on C on the other hand.* In the real domain this problem has been extensively studied;[4] the complex domain has received less attention.[5]

Received February 21, 1940; presented to the American Mathematical Society, September, 1936.

[1] A function which can be expressed in the form $a_0z^n + a_1z^{n-1} + \cdots + a_n$ is called a polynomial in z of degree n; we do not assume $a_0 \neq 0$.

[2] See for instance Walsh [1]; here and elsewhere numbers in square brackets refer to the bibliography at the end of this paper.

The reader is referred to that same work (Walsh [1]) for terminology not explicitly defined in the present paper.

[3] See for instance Walsh [1], pp. 65 ff.

[4] See for instance Bernstein [1, 2]; de la Vallée Poussin [1]; Jackson [1].

[5] See Sewell [1, 2, 3]; Curtiss [1, 2]. Jackson [2, 3] deduces the order of Tchebycheff approximation from approximation as measured by an integral.

Let $f(z)$ be analytic throughout the interior of some C_ρ; Problem β is the study of the relation between degree of convergence of $p_n(z)$ to $f(z)$ on E on the one hand, and continuity properties of $f(z)$ on C_ρ on the other hand. In the case where $f(z)$ has poles or essential singularities on C_ρ, Problem β has been studied by various writers;[6] if the function $f(z)$ has a bounded generalized derivative on C_ρ, or if some derivative of $f(z)$ satisfies a Lipschitz condition on C_ρ, Problem β has been studied less extensively.[7] The main purpose of the present paper is to study Problem β, when $f(z)$ or one of its derivatives is assumed to satisfy a Lipschitz condition of order α $(0 < \alpha \leqq 1)$ on C_ρ; we establish results concerning degree of convergence on E.

An additional problem which we shall call *Problem γ is the study of the degree of convergence on C_ρ and on $C_{\rho'}$ $(1 < \rho' < \rho)$ (but not on E) of polynomials arising out of the study of Problem β.* As far as the present writers are aware, there are no previous results in the literature on this problem, except perhaps in the special case of Taylor's series.

We remark here that results on approximation to analytic functions by polynomials in the complex variable are analogous to, and have application to, results on approximation to harmonic functions by harmonic polynomials. The writers hope shortly to indicate in detail this relationship for Problems α, β, and γ.

We shall now be more explicit about the new results of the present paper.

2. **Summary.** Suppose $f(z)$ is analytic interior to C_ρ, continuous in the corresponding closed region or regions. In Chapter I (§3) we establish the formula

$$(1) \qquad f(z) - p_n(z) = \frac{1}{2\pi i}\int_{C_\rho} \frac{\omega_n(z)[f(t) - P_n(t)]\,dt}{\omega_n(t)(t-z)}, \qquad z \text{ interior to } C_\rho,$$

where $\omega_n(z) \equiv (z - z_1)(z - z_2) \cdots (z - z_{n+1})$, where the z_j are arbitrary points interior to C_ρ, where $P_n(z)$ is an arbitrary polynomial of degree n, and where $p_n(z)$ is the unique polynomial of degree n which interpolates to $f(z)$ in the points z_j. Equation (1) enables us to employ known results on degree of approximation of polynomials $P_n(z)$ to $f(z)$ on C_ρ (i.e., Problem α) to obtain new results on degree of approximation of polynomials $p_n(z)$ to $f(z)$ on E (i.e., Problem β).

Chapter II (§§4–6) is devoted to a study of asymptotic properties of polynomials whose roots are assigned points of interpolation; the latter are chosen in various cases as equally distributed points, roots of the Faber, Tchebycheff, and orthogonal polynomials, and the poles of lemniscates. Our primary object in Chapter II is to show that the function $\rho^n\omega_n(z)/\omega_n(t)$ [or $(\rho^{2n}/|\,\omega_n(t)\,|^2)\cdot\int_C |\,\omega_n(z)\,|^2\,|\,dz\,|$ for the orthogonal polynomials] is uniformly

[6] Especially de la Vallée Poussin [1], Faber [2], Bernstein [1, 2], Sewell [3].

[7] Sewell [3], Walsh and Sewell [2, 3, 4].

bounded for z on E and t on C_ρ, where the z_i are the points of interpolation mentioned.

Chapter III (§§7–9) contains the main results of the paper, namely, the actual application of formula (1) and its corollaries by the use of the asymptotic relations of Chapter II. We thereby establish quite definitive results on Problem β when $f(z)$ or one of its derivatives is assumed to satisfy a generalized Lipschitz condition on C_ρ.

The general question of the relation between Lipschitz conditions on C_ρ and Lipschitz conditions in the corresponding closed regions is considered in §10. In the latter parts (§§11, 12) of the concluding Chapter IV (§§10–12), we consider (§11) Problem γ, the behavior on $C_{\rho'}$ of the various polynomials defined in Chapter III. In §12 we finally obtain inequalities on the error in the case of best approximation, and thereby prove that the results of Chapter III are in a sense the best possible.

Chapter I. General inequalities on polynomial approximation

3. Fundamental formula and algebraic inequalities. Our primary object in this section is the proof of

THEOREM 3.1. *Let E be a closed limited point set whose boundary is denoted by C, and whose complement K is connected and regular. Let C_ρ, $\rho > 1$, be a level locus of Green's function for K with pole at infinity. Let $f(z)$ be analytic interior to C_ρ and represented interior to C_ρ by the Cauchy integral over C_ρ. Let the points $z_1, z_2, \cdots, z_{n+1}$ lie interior to C_ρ, and let us set*

$$(3.1) \qquad \omega_n(z) \equiv (z - z_1)(z - z_2) \cdots (z - z_{n+1}).$$

Then we have

$$(3.2) \qquad f(z) - p_n(z) \equiv \frac{1}{2\pi i} \int_{C_\rho} \frac{\omega_n(z)[f(t) - P_n(t)]\,dt}{\omega_n(t)(t - z)}, \qquad z \text{ interior to } C_\rho,$$

where $P_n(z)$ is an arbitrary polynomial of degree n, and where $p_n(z)$ is the unique polynomial of degree n which interpolates to $f(z)$ in the points z_i.

Here C_ρ may consist of several rectifiable Jordan curves, mutually exterior except perhaps for a finite number of points.[8]

Under the hypothesis of Theorem 3.1 we have the Lagrange-Hermite interpolation formula (e.g., Montel [1] or Walsh [1]):

$$(3.3) \qquad f(z) - p_n(z) \equiv \frac{1}{2\pi i} \int_{C_\rho} \frac{\omega_n(z) f(t)\,dt}{\omega_n(t)(t - z)}, \qquad z \text{ interior to } C_\rho.$$

In particular we may set $f(z) \equiv P_n(z)$ in (3.3); then the two polynomials $p_n(z)$

[8] Of course formula (3.2) holds also even if the integral is taken not over C_ρ but over a much more general curve or set of curves, when $f(z)$ is assumed represented by the corresponding Cauchy integral. The more general formula has no application in the present paper.

and $P_n(z)$ each of degree n coincide in the $n + 1$ points z_j and are identically equal, so we have[9]

$$(3.4) \qquad 0 \equiv \frac{1}{2\pi i}\int_{C_\rho} \frac{\omega_n(z)P_n(t)\,dt}{\omega_n(t)(t - z)}, \qquad z \text{ interior to } C_\rho.$$

If we subtract (3.4) from (3.3) in its original form [i.e., without particularizing $f(z)$], we obtain (3.2) and the theorem.

By inspection of (3.2) we now have the useful inequality

$$(3.5) \quad |f(z) - p_n(z)| \leqq \frac{1}{2\pi}\int_{C_\rho} \frac{|\omega_n(z)|\cdot|f(t) - P_n(t)|\cdot|dt|}{|\omega_n(t)|\cdot|t - z|}, \qquad z \text{ interior to } C_\rho.$$

THEOREM 3.2. *Under the hypothesis of Theorem* 3.1, *the inequalities*

$$(3.6) \qquad |f(z) - P_n(z)| \leqq \epsilon_n, \qquad z \text{ on } C_\rho,$$

$$(3.7) \qquad \left|\frac{\omega_n(z)}{\omega_n(t)}\right| \leqq \frac{M}{\rho^n}, \qquad z \text{ on } E \text{ and } t \text{ on } C_\rho,$$

where M is a constant independent of n, t, and z, imply the inequality

$$|f(z) - p_n(z)| \leqq \frac{M_1\epsilon_n}{\rho^n}, \qquad z \text{ on } E,$$

where M_1 is a constant independent of n and z.

By means of the Hölder inequality

$$(3.8) \qquad \left|\int F^\alpha G^{1-\alpha}\right| \leqq \left(\int |F|\right)^\alpha\left(\int |G|\right)^{1-\alpha}, \qquad 0 < \alpha < 1,$$

we find from (3.2)

$$|f(z) - p_n(z)|$$

$$(3.9) \quad \leqq \frac{1}{2\pi}\left(\int_{C_\rho}\left|\frac{\omega_n(z)}{\omega_n(t)(t - z)}\right|^{1/\alpha}\cdot|dt|\right)^\alpha\left(\int_{C_\rho}|f(t) - P_n(t)|^{1/(1-\alpha)}|dt|\right)^{1-\alpha},$$

$$0 < \alpha < 1, \; z \text{ interior to } C_\rho.$$

Inequality (3.9) *is important in that it permits us to use a known degree of approximation by polynomials $P_n(z)$ to $f(z)$ as measured by an integral over C_ρ (for instance, polynomials $P_n(z)$ of best approximation on C_ρ in the sense of least p-th powers) to obtain results on Tchebycheff approximation to $f(z)$ on E:*

THEOREM 3.3. *Under the hypothesis of Theorem* 3.1, *the inequalities* (3.7) *and*

$$(3.10) \qquad \int_{C_\rho} |f(t) - P_n(t)|^{1/(1-\alpha)}\cdot|dt| \leqq \epsilon_n, \qquad 0 < \alpha < 1,$$

[9] Equation (3.4) exhibits an orthogonality relation, perhaps of some interest in itself, between the polynomial $P_n(t)$ and the rational function $1/[\omega_n(t)(t - z)]$, where z is an arbitrary point interior to C_ρ. Of course (3.4) admits of direct proof.

imply the inequality

$$|f(z) - p_n(z)| \leqq \frac{M_1 \epsilon_n^{1-\alpha}}{\rho^n}, \qquad z \text{ on } E,$$

where M_1 is a constant independent of n and z.

By taking conjugates in (3.2) we have by use of (3.2) itself

$$\int_C |f(z) - p_n(z)|^2 |dz| = \frac{1}{4\pi^2} \int_C |\omega_n(z)|^2 \cdot \left| \int_{C_\rho} \frac{[f(t) - P_n(t)]\, dt}{\omega_n(t)(t - z)} \right|^2 \cdot |dz|.$$

This formula gives an evaluation of the error in the sense of least squares on C, an evaluation involving $\omega_n(z)$ on C merely through $\int_C |\omega_n(z)|^2 |dz|$.

THEOREM 3.4. *Under the hypothesis of Theorem 3.1, the inequalities*

$$|f(z) - P_n(z)| \leqq \epsilon_n, \qquad z \text{ on } C_\rho,$$

$$\int_C |\omega_n(z)|^2 |dz| \leqq \delta_n, \qquad |\omega_n(t)| \geqq \mu_n \quad \text{for } t \text{ on } C_\rho,$$

imply the inequality

$$\int_C |f(z) - p_n(z)|^2 |dz| \leqq \frac{M_1 \delta_n \epsilon_n^2}{\mu_n^2},$$

where M_1 is a constant independent of n.

The significance of Theorems 3.2, 3.3, and 3.4 lies in (i) showing the existence of sequences of polynomials $p_n(z)$ which converge to $f(z)$ on C with a certain degree of convergence, and (ii) studying the degree of convergence of the explicit polynomials $p_n(z)$ found by interpolation to $f(z)$ in variously chosen specific sets of points z_j.

Chapter II. Roots of polynomials as points of interpolation

4. **Equally distributed points.** Let the boundary C of E consist of the mutually exterior rectifiable Jordan curves $C_1, C_2, \cdots, C_\nu$, assumed to be smooth in a sense later to be made more precise. In the notation of §1 we have $G(x, y) = \log (x^2 + y^2)^{\frac{1}{2}} + G_0(x, y)$, where $G_0(x, y)$ approaches a constant value $-g$ as (x, y) becomes infinite. If we set $V(x, y) = G(x, y) + g$, we have[10]

$$V(x, y) = \int_C \varphi_1(s) \log r\, ds, \qquad \varphi_1(s) = \frac{1}{2\pi} \frac{\partial V}{\partial n}, \tag{4.1}$$

where n is the exterior normal for C, where $r = |z - \zeta|$, $ds = |d\zeta|$, and $z = x + iy$ is an arbitrary point of K. Indeed, we shall prove

LEMMA 4.1. *Suppose C has a tangent at every point, suppose $\partial G/\partial n$ is uniformly bounded and measurable on C, and suppose there exists a constant Q such that*

[10] See for instance Walsh [1], pp. 68 ff., an exposition which is followed closely in this section. The method is due to Hilbert, Fejér, Szegö, and others. Compare also Montel [1].

$$(4.2) \qquad \left| \int_{\zeta_1}^{\zeta_2} ds \right| \leqq Q |\zeta_2 - \zeta_1|,$$

where the integral is taken along C, and where Q is independent of ζ_1 and ζ_2; the integral in (4.2) *signifies the integral over the shorter of the two paths over C connecting ζ_1 and ζ_2, where ζ_1 and ζ_2 lie on the same C_j. Then equation* (4.1) *is valid also for z on C.*

We assume $|\partial G/\partial n| \leqq M$ on C. Let $z_0 = x_0 + iy_0$ be fixed on C. Since $V(x, y)$, when suitably defined on C, is known to be continuous in the closed region $K + C$, it is sufficient for us to prove that

$$(4.3) \qquad \lim_{z_1 \to z_0} \int_C \frac{\partial V}{\partial n} \log \left| \frac{z_1 - \zeta}{z_0 - \zeta} \right| ds = 0,$$

where z_1 is in K and lies on the normal to C at z_0. The existence of the integral in (4.3) will appear in the course of the proof.

Let $\eta > 0$ be arbitrary, but so small that we have $\eta < 1$ and $\log (1 - \eta^2) > -\eta$. From this it follows that we have also $\log (1 + \eta^2) < \eta$. Let C_1' be the set of points ζ on C such that $|\zeta - z_0| \geqq \eta$, and let C_2' be the set $C - C_1'$. We shall choose z_1 in such a way that we have $|z_1 - z_0| < \eta^3$, whence it follows that we have for ζ on C_1'

$$\left| \frac{z_1 - \zeta}{z_0 - \zeta} \right| \leqq \frac{|z_1 - z_0| + |z_0 - \zeta|}{|z_0 - \zeta|} \leqq 1 + \eta^2,$$

$$\left| \frac{z_1 - \zeta}{z_0 - \zeta} \right| \geqq \frac{|z_0 - \zeta| - |z_1 - z_0|}{|z_0 - \zeta|} \geqq 1 - \eta^2,$$

$$\left| \log \left| \frac{z_1 - \zeta}{z_0 - \zeta} \right| \right| < \eta, \qquad \zeta \text{ on } C_1',$$

$$(4.4) \qquad \left| \int_{C_1'} \frac{\partial V}{\partial n} \log \left| \frac{z_1 - \zeta}{z_0 - \zeta} \right| ds \right| \leqq Ml\eta,$$

where l is the length of C.

For ζ on C_2' we have

$$|\zeta - z_1| \leqq |\zeta - z_0| + |z_0 - z_1| \leqq \eta + \eta^3,$$

$$\log \left| \frac{\zeta - z_1}{\zeta - z_0} \right| = \log |\zeta - z_1| - \log |\zeta - z_0|$$

$$\leqq \log (\eta + \eta^3) - \log |\zeta - z_0|.$$

Let C_2' lie in a closed double sector of angle $2\alpha < \pi$ with vertex z_0; such a sector exists by virtue of the existence of a tangent to C at z_0. Since z_1 lies on the normal to C at z_0, we may assume (by a possible further restriction on η) that z_1 lies on the bisector of the exterior angle of this sector. Since the distance

from ζ on C_2' to this bisector is no greater than $|\zeta - z_1|$, and is at least as great as $|\zeta - z_0| \cos \alpha$, we have

$$|\zeta - z_1| \geqq |\zeta - z_0| \cos \alpha,$$

$$\log \left| \frac{\zeta - z_1}{\zeta - z_0} \right| \geqq \log \cos \alpha, \qquad \zeta \text{ on } C_2'.$$

In sum, we may write for ζ on C_2'

$$\left| \log \left| \frac{\zeta - z_1}{\zeta - z_0} \right| \right| \leqq -\log \cos \alpha + \log (\eta + \eta^3) - \log |\zeta - z_0|.$$

By (4.2) the path of integration is of length not greater than $2Q\eta$, whence

$$\begin{aligned} \left| \int_{C_2'} \frac{\partial V}{\partial n} \log \left| \frac{z_1 - \zeta}{z_0 - \zeta} \right| ds \right| &\leqq -2MQ\eta \log \cos \alpha + 2MQ\eta \log \eta \\ &\quad + 2MQ\eta^2 - 2MQ \int_{z_0}^{z_0+\eta} \log |\zeta - z_0| \cdot |d\zeta|, \end{aligned} \tag{4.5}$$

where the last integral is interpreted as an ordinary definite integral rather than a line integral; inequality (4.5) implies the existence of the second member of (4.1) for z on C. The right member of (4.5) may be written

$$-2MQ\eta \log \cos \alpha + 2MQ\eta \log \eta + 2MQ\eta^2 + 2MQ[\eta - \eta \log \eta], \tag{4.6}$$

which by suitable choice of η can be made as small as desired.

That is to say, if $\epsilon > 0$ is given, we choose η with $0 < \eta < 1$ so that we have $\log (1 - \eta^2) > -\eta$, so that (4.6) plus $Ml\eta$ is less than ϵ, and so that the double sector with vertex z_0 has the properties prescribed. For z_1 in K on the normal to C at z_0 and with $|z_1 - z_0| < \eta^3$, the first member of (4.5) plus the first member of (4.4) is less than ϵ, so (4.3) is established.

In equation (4.1) we assume that $\partial V/\partial n$ exists, and is uniformly bounded and measurable on C. We set

$$u_0 = \int_C \varphi_1(s)\, ds, \qquad u(\zeta) = \int_0^{s(\zeta)} \varphi_1(s)\, ds. \tag{4.7}$$

For definiteness and simplicity in exposition we henceforth frequently assume $\nu = 2$; the extension to arbitrary finite ν is immediate and is left to the reader. By the fundamental properties of the function u we know (op. cit., p. 71)

$$u_0 = \int_C du = 1. \tag{4.8}$$

The boundary C of E consists of the Jordan curves C_1 and C_2, and we set

$$\int_{C_1} du = \lambda, \qquad \int_{C_2} du = 1 - \lambda, \qquad 0 < \lambda < 1.$$

Of course the function $w = \varphi(z)$ which maps K onto the region $|w| > 1$ is not

uniquely determined, and even when chosen is multiple valued; but if we choose a particular branch, the map sets up a correspondence between C_1 and an arc of length $2\pi\lambda$ of the circle $\gamma: |w| = 1$; we may choose the function $\varphi(z)$ and the branch so as to measure the arc of length $2\pi\lambda$ from any preassigned point of γ, corresponding to any preassigned point of C_1. Similarly, we obtain a correspondence with similar properties between C_2 and an arbitrary arc of γ of length $2\pi(1 - \lambda)$; each of these arcs is to be taken as closed at one end, open at the other. Thus we obtain a one-to-one correspondence between the points of C and the points of γ, a correspondence which is continuous on C except at a single point of C_1 and a single point of C_2, and which is continuous on γ except at the points of γ corresponding to those two points of C_1 and C_2, respectively. This correspondence, once chosen, is to be temporarily fixed during our subsequent discussion; we shall eventually need another similar correspondence. Certain continuity and differentiability properties of the correspondence can be obtained from the continuity and differentiability properties of C alone. We are interested in these properties merely in the neighborhood of C and γ; we do not need continuity properties in regions remote from C and γ.

Equations (4.1), (4.7), and (4.8) yield

$$V(x, y) = \int_0^1 \log r \cdot du, \qquad z \text{ in } K,$$

so by definition of the definite integral we have for z in K

$$V(x, y) = \lim_{n \to \infty} \frac{1}{n} (\log r_1 + \log r_2 + \cdots + \log r_n),$$

where $r_1, r_2, \cdots, r_n$ are the distances from z to the n points ζ_m of C which correspond to the n equidistant values k/n $(k = 1, 2, \cdots, n)$ of u in the interval $(0, 1)$; of course the points ζ_m depend on n, but for simplicity that dependence is not indicated in the notation. Otherwise expressed, the points ζ_m may be considered as the transforms of the roots of unity on γ, and the points ζ_m are said to be *equally distributed*[11] on C.

Let us now write

$$\omega_{n-1}(z) \equiv (z - \zeta_1)(z - \zeta_2) \cdots (z - \zeta_n);$$

we wish to find an asymptotic expression for this function, for z on C or in K. Let $z = \psi(w)$ be the inverse of $w = \varphi(z)$, and let us set $\zeta_m = \psi(w_m)$, $w_m = e^{2\pi i m/n}$. We introduce the notation

$$(4.9) \quad \tau_n = \frac{\omega_{n-1}(z)}{w^n - 1} = \frac{[\psi(w) - \psi(w_1)][\psi(w) - \psi(w_2)] \cdots [\psi(w) - \psi(w_n)]}{(w - w_1)(w - w_2) \cdots (w - w_n)}.$$

Equation (4.9) suggests a convenient continuity condition:

DEFINITION 4.2. The set C of mutually exterior Jordan curves $C_1, C_2, \cdots, C_\nu$

[11] See for instance Walsh [1], pp. 70 ff. The term is due to Fejér.

will be called a *contour* provided each C_j has a tangent at every point; provided the function

$$\log \left| \frac{\psi(w) - \psi(w')}{w - w'} \right| \tag{4.10}$$

for each possible definition of $\psi(w)$ and $\psi(w')$ is locally bounded in the two-dimensional sense for $|w| = 1$ and $|w'| \geqq 1$ in the neighborhood of each point of continuity of $\psi(w)$; and provided the function (4.10) satisfies on each arc γ_j of $\gamma: |w| = 1$ corresponding to a curve C_j a Lipschitz condition[12] of order 1 in w, uniformly with respect to w' on any closed arc of $|w'| = 1$ interior to a γ_k.

On each arc γ_j of γ corresponding to a curve C_j a single branch of $\psi(w)$ is to be used here.

The interpretation and significance of this definition require some discussion. Each arc γ_j is closed at one end, open at the other end; such end-points are points of discontinuity of $\psi(w)$, for on one side of such a point the value $z = \psi(w)$ lies on one C_j and on the other the value $z = \psi(w)$ lies on a different C_j. This unusual and purely accidental character disappears if interpreted geometrically, thanks to our requirement that the Lipschitz condition shall be satisfied no matter how $\psi(w)$ is defined, that is to say, no matter which point of each C_j is chosen to correspond to an end-point of γ_j, and no matter where the latter lies on γ.

The Lipschitz condition demanded may be written

$$\left| \log \left| \frac{\psi(w_1) - \psi(w')}{w_1 - w'} \right| - \log \left| \frac{\psi(w_2) - \psi(w')}{w_2 - w'} \right| \right| \leqq A |w_1 - w_2|, \tag{4.11}$$

where A is independent of w_1, w_2, and w'. Choose w_1 as that end-point of γ_1 which belongs to γ_1, choose w_2 as a fixed interior point of γ_1, and choose w' as a variable point of the other arc γ_j having w_1 as an end-point. When w' approaches w_1, the distance $|\psi(w_1) - \psi(w')|$ is the distance from a point of C_1 to a point of C_j, hence is bounded from zero, so the expression

$$\left| \frac{\psi(w_1) - \psi(w')}{w_1 - w'} \right|$$

becomes infinite. When w' approaches w_1, the expression

$$\left| \frac{\psi(w_2) - \psi(w')}{w_2 - w'} \right|$$

is bounded, so the left member of (4.11) becomes infinite, and (4.11) cannot hold uniformly with respect to w' with w' merely restricted by the equation $|w'| = 1$; this is our justification for requiring in Definition 4.2 uniformity with respect

[12] A function $f(z)$ defined on a set E is said to satisfy a Lipschitz condition of order α $(0 < \alpha \leqq 1)$ if there exists M such that $|f(z_1) - f(z_2)| \leqq M |z_1 - z_2|^\alpha$ whenever z_1 and z_2 lie on E. Unless otherwise stated, a Lipschitz condition is considered to be of order unity: $\alpha = 1$.

to w' merely for w' on any closed arc of $|w'| = 1$ interior to a γ_k. But we suppose here too that this condition holds for every choice of a point of C_k corresponding to an end-point of γ_k, that is to say, for an arbitrary choice of the correspondence between γ and C.

Strictly speaking, the function (4.10) is not defined if $w = w'$. However, let us assume (4.11) valid, where now w' as well as w_1 and w_2 lies interior to a particular arc γ_j. Denote by $W_1(w_1)$ any limiting value of the function

$$\log \left| \frac{\psi(w_1) - \psi(w')}{w_1 - w'} \right|$$

as w' approaches $w_1 \neq w_2$. By the uniformity of (4.11) and the continuity of $\psi(w)$ for $w = w_1$ it follows that $W_1(w_1)$ is finite and

$$\left| W_1(w_1) - \log \left| \frac{\psi(w_2) - \psi(w_1)}{w_2 - w_1} \right| \right| \leqq A |w_1 - w_2|. \tag{4.12}$$

If we allow w_2 to approach w_1, it follows from (4.12) that

$$\log \left| \frac{\psi(w_2) - \psi(w_1)}{w_2 - w_1} \right|$$

approaches the finite limit $W_1(w_1)$. Hence $W_1(w_1)$ is uniquely determined. By a change of notation in (4.12) we have

$$\left| W_1(w_2) - \log \left| \frac{\psi(w_1) - \psi(w_2)}{w_1 - w_2} \right| \right| \leqq A |w_1 - w_2|,$$

whence, by (4.12) itself,

$$| W_1(w_1) - W_1(w_2) | \leqq 2A \, | w_1 - w_2 |, \tag{4.13}$$

a Lipschitz condition of order unity for the function $W_1(w)$ on any arc γ_j.

When w' approaches w_1, both points lying on the same γ_j, it follows from the fact that C_j has a tangent that

$$\arg \left[\frac{\psi(w_1) - \psi(w')}{w_1 - w'} \right]$$

approaches a limit. Consequently when w' approaches w_1, the function

$$\log \left[\frac{\psi(w_1) - \psi(w')}{w_1 - w'} \right] = \log \left| \frac{\psi(w_1) - \psi(w')}{w_1 - w'} \right| + i \arg \left[\frac{\psi(w_1) - \psi(w')}{w_1 - w'} \right] \tag{4.14}$$

approaches a limit, properly denoted by $\log \psi'(w_1)$, where the derivative $\psi'(w_1)$ is defined in the one-dimensional sense on γ_j and is different from zero.

The function

$$\frac{\psi(w_1) - \psi(w')}{w_1 - w'}$$

considered as a function of w', with $|w'| \geqq 1$ and w_1 a fixed interior point of γ_j, is continuous in the two-dimensional sense on any closed arc of $|w'| = 1$

interior to γ_j except perhaps in the point $w' = w_1$, it approaches the same limit when w' on γ_j approaches w_1 in either sense, and by Definition 4.2 it is bounded in the two-dimensional sense in the neighborhood of w_1. It follows from a theorem due to Lindelöf that the limit in the two-dimensional sense exists as w' exterior to γ approaches w_1, and is $\psi'(w_1) \neq 0$; the derivative $\psi'(w_1)$ is now a derivative in the two-dimensional sense.

The first member of (4.14), considered as a function of w', is continuous and its real part satisfies a Lipschitz condition of order unity on any closed arc of $|w'| = 1$ interior to γ_j. Hence by a theorem due to Privaloff [1][13] the pure imaginary part of that function (and therefore also the function itself) satisfies on any such arc a Lipschitz condition of every order α $(0 < \alpha < 1)$. It follows by the reasoning used for inequality (4.11) in the proof of (4.13) that we have on the arc considered

$$(4.15) \qquad |\log \psi'(w_1) - \log \psi'(w_2)| \leqq A_1 |w_1 - w_2|^\alpha.$$

In the further study of $\psi'(w)$ we shall use

LEMMA 4.3. *Let* $\log \Phi(w)$ *satisfy a Lipschitz condition of given order* α $(0 < \alpha \leqq 1)$ *on a closed proper subarc of* $|w| = 1$. *Then on that arc the function* $\Phi(w)$ *also satisfies such a condition.*

Our hypothesis involves the boundedness of $\log \Phi(w)$, a consequence of the inequality

$$|\log \Phi(w_1) - \log \Phi(w_2)| \leqq B_1 |w_1 - w_2|^\alpha.$$

The continuity of $\Phi(w)$ on the arc considered is a consequence of the continuity of $\log \Phi(w)$ there. For $w_1 - w_2$ sufficiently small, namely, $|\Phi(w_1) - \Phi(w_2)| < \lambda \cdot \min |\Phi(w)|$ $(0 < \lambda < 1)$, we have

$$\left|\log \frac{\Phi(w_1)}{\Phi(w_2)}\right| = \frac{|\Phi(w_1) - \Phi(w_2)|}{|\Phi(w_2)|} \left| 1 - \frac{1}{2}\left(\frac{\Phi(w_1)}{\Phi(w_2)} - 1\right) + \cdots \right|.$$

[13] The theorem of Privaloff is ordinarily stated for the conjugate $v(z)$ of a function $u(z)$ which is harmonic for $|z| < 1$, continuous for $|z| \leqq 1$, with a Lipschitz condition of order unity *on the entire circumference* $|z| = 1$; but we shall now prove that *if $u(z)$ is harmonic in the deleted neighborhood of a closed arc A of that circumference, continuous on A in the two-dimensional sense, and satisfies a Lipschitz condition of order unity on A, then the conjugate $v(z)$ defined on A so as to be continuous there satisfies on any closed proper subarc of A a Lipschitz condition of every order α $(0 < \alpha < 1)$.* Let $u_1(z)$ be defined on $|z| = 1$ equal to $u(z)$ on A, and satisfying a Lipschitz condition over the whole circumference (it is sufficient to take $u_1(z)$ exterior to A as a linear function of arc length), and then let $u_1(z)$ defined by these boundary values be continuous for $|z| \leqq 1$, harmonic for $|z| < 1$. By the original form of Privaloff's theorem, the conjugate $v_1(z)$ satisfies a Lipschitz condition of order α on $|z| = 1$. The function $u(z) - u_1(z)$ is harmonic in the deleted neighborhood of A, continuous on A and equal to zero there, and can be extended harmonically across A. The conjugate $v(z) - v_1(z)$ satisfies a Lipschitz condition of order α on any closed proper subarc of A, hence so also does $v(z)$ itself.

This last factor is of the form

$$\left| \frac{\log \zeta}{\zeta - 1} \right|, \qquad | \zeta - 1 | < \lambda < 1,$$

and hence is bounded from zero, whence by virtue of our hypothesis we have

$$| \Phi(w_1) - \Phi(w_2) | \leqq B \, | w_1 - w_2 |^\alpha \tag{4.16}$$

for $w_1 - w_2$ sufficiently small. By the general inequality

$$| \xi_1 + \xi_2 |^p \leqq 2^{p-1}[| \xi_1 |^p + | \xi_2 |^p], \qquad p \geqq 1,$$

and by use of (4.16) for two pairs of points (w_1, w_2) and (w_2, w_3), with w_2 between w_1 and w_3, we may write

$$\begin{aligned} | \Phi(w_1) - \Phi(w_3) |^{1/\alpha} &\leqq 2^{(1-\alpha)/\alpha}[| \Phi(w_1) - \Phi(w_2) |^{1/\alpha} \\ &+ | \Phi(w_2) - \Phi(w_3) |^{1/\alpha}] \leqq 2^{(1-\alpha)/\alpha} B^{1/\alpha}[| w_1 - w_2 | + | w_2 - w_3 |]. \end{aligned} \tag{4.17}$$

No matter what proper subarc of $| w | = 1$ is considered, the ratio of the arc $w_1w_2w_3$ to the chord w_1w_3 has a positive upper bound b which is independent of the points w_1, w_2, w_3; we have

$$| w_1 - w_2 | + | w_2 - w_3 | \leqq b \, | w_1 - w_3 |,$$

whence by (4.17)

$$| \Phi(w_1) - \Phi(w_3) | \leqq B' \, | w_1 - w_3 |^\alpha,$$

where B' does not depend on w_1, w_2, and w_3. By virtue of the uniform continuity of the function $\Phi(w)$ on the arc of $| w | = 1$ considered, it is now clear that use of (4.16) for a suitably chosen finite number of pairs of points w_1 and w_2 yields the conclusion of Lemma 4.3.

The restriction in Lemma 4.3 that the closed arc considered should be a proper subarc of $| w | = 1$ is unnecessary, for if the arc is the entire circumference, the proof just given requires only obvious modifications.

A further preliminary result is

LEMMA 4.4. *If C is a contour, the function $\psi'(w)$ exists and is continuous in the two-dimensional sense on any closed arc interior to a γ_j, is different from zero on such an arc, and satisfies there a Lipschitz condition of every order α $(0 < \alpha < 1)$.*

We have already shown the existence in the two-dimensional sense of the derivative $\psi'(w)$ on the arc considered; the continuity in the one-dimensional sense of $\psi'(w)$ and Lipschitz condition of order α there are a consequence of (4.15) and Lemma 4.3; it follows from (4.15) that $\psi'(w) \neq 0$ on γ_j.

The continuity of $\psi'(w)$ in the two-dimensional sense remains to be established. For z on C we write $ds = | dz |$, $z = \psi(w)$,

$$\frac{dz}{ds} = \frac{\psi'(w)}{| \psi'(w) |} \frac{dw}{| dw |}. \tag{4.17a}$$

The functions $\psi'(w)$, $1/|\psi'(w)|$, $dw/|dw|$ all satisfy Lipschitz conditions of some order with respect to w, so dz/ds satisfies such a condition. We have also

$$|w_2 - w_1| \leq \int_{w_1}^{w_2} |dw| = \int_{s_1}^{s_2} \frac{ds}{|\psi'(w)|} \leq K_1 |s_2 - s_1|,$$

where K_1 is suitably chosen and is independent of w_1 and w_2, and where $w_1 = w(s_1)$, $w_2 = w(s_2)$. Consequently $z(s)$ has the property that $z'(s)$ exists and satisfies a Lipschitz condition of some order *with respect to s*. This property holds on every arc of C. It now follows from an important theorem due to Kellogg [1] that $\psi'(w)$ exists and is continuous in the two-dimensional sense on $|w| = 1$. This completes the proof of Lemma 4.4.

It follows from Lemma 4.4 that if C is a contour, $\partial V/\partial n$ is uniformly bounded and measurable on C. Moreover, if C is a contour, we have on C the equations $z = \psi(w)$ and (4.17a), so arc and chord are equivalent infinitesimals, and (4.2) is valid for suitably chosen Q. Thus the conditions of Lemma 4.1 are fulfilled.

For reference we now state another lemma (Walsh and Sewell [1]).

LEMMA 4.5. *Let $f(x)$ be continuous in the closed interval $0 \leq x \leq a$ except for a finite number of finite discontinuities. Let $f(x)$ satisfy one and the same Lipschitz condition of order unity in every subinterval of continuity. Then we have*

$$\left| \int_0^a f(x)\,dx - \frac{a}{n} \sum_{k=1}^{n} f(ak/n) \right| \leq \frac{M}{n},$$

where M is a suitably chosen constant depending on the constant in the Lipschitz condition and on the sum of the discontinuities of $f(x)$ but independent of n.

Let now C be a contour. With the notation (4.7), equation (4.1) may be written for z on C (see Lemma 4.1)

$$\int_C \log r \cdot du = V(x, y) = g, \tag{4.18}$$

where $r = |z - \zeta|$, and where $\zeta = \zeta(u)$ is the running variable on C for the integration. Equation (4.18) may also be written in the form

$$\int_{|w'|=1} \log |\psi(w) - \psi(w')| \cdot du = g, \tag{4.18$'$}$$

where w lies on the circle $|w| = 1$. *These two equations are to be considered identical.* A consequence of this convention is that in (4.18′) we assume a one-to-one correspondence between the points of $|w'| = 1$ and the points of C; but this correspondence is, as we have already said, not uniquely determined. The function $\psi(w)$ itself is not uniquely determined, and once it has been determined is not single valued (perhaps infinitely many valued) on $|w| = 1$. We do not assume any necessary connection between the various branches or determinations of $\psi(w)$ employed on the various arcs γ_i, provided that the cor-

respondence between points of $| w' | = 1$ and points of C is continuous on each such arc and in the large is one-to-one.

In equations (4.18) and (4.18′) we have $w' = \varphi(z) = e^{V+iV'-g}$, where V' is conjugate to V in K, whence for z on C and $| w' | = 1$, we have $dw' = w'(dV + idV')$, $| dw' | = dV' = (\partial V/\partial n)\, ds = 2\pi\, du$.

A special case of equation (4.18), valid by Lemma 4.1, occurs when C is the circle $| w | = 1$:

$$\frac{1}{2\pi}\int_{|w'|=1} \log | w - w' | \cdot | dw' | = 0, \tag{4.19}$$

for w on the circle $| w | = 1$. By virtue of the map of the original region K onto $| w | > 1$, equations (4.18) and (4.19) can be combined into the form

$$\int_{|w'|=1} \log \left| \frac{\psi(w) - \psi(w')}{w - w'} \right| \cdot du = g, \tag{4.20}$$

where w lies on the circle $| w | = 1$.

Since C is a contour, the integrand in (4.20) satisfies on each γ_j a Lipschitz condition with respect to u or w', uniformly for w on each closed arc interior to a γ_k. Equations (4.20) and (4.9) yield by Lemma 4.5 for $| w | = 1$ or z on C

$$\left| g - \frac{1}{n} \log | \tau_n | \right| \leqq \frac{M}{n}, \qquad | ng - \log | \tau_n | | \leqq M;$$

it follows from the proof of Lemma 4.5 that M can be chosen as independent of both n and w, where w lies on an arbitrary closed arc interior to a γ_j, or in other words when z lies on any closed proper subarc of any C_j. Consequently, *by use of a new correspondence* between γ and C, involving new points of C_j as images of the end-points of the γ_j, it follows that these inequalities are valid for suitable choice of M, uniformly for *all* z on C_j.

When such a new correspondence is set up between γ and C, care must be taken not to alter the *geometric* significance of the points ζ_m on C. This geometric significance is to be fixed once for all by the original correspondence.[14] The constant M which occurs in Lemma 4.5 involves merely (a) the constant in the assumed Lipschitz condition, and (b) the sum of the magnitude of the finite jumps of the function integrated; but M does not depend on the value of the numbers b and c, if $f(x)$ is supposed periodic with period a and if we replace the conclusion of Lemma 4.5 by

$$\left| \int_b^c f(x)\, dx - \frac{a}{n} \sum f(ak/n) \right| \leqq \frac{M}{n},$$

where the summation is extended over the N points ak/n in the interval $b \leqq x < c$, with $c - b < a$. The new correspondence, say between γ_1 and C_1,

[14] For definiteness we chose the points $w_m = \varphi(\zeta_m)$ as the n-th roots of unity; they may equally well be chosen as the n-th roots of a number β_n of modulus unity but which varies with n.

is best broken into a correspondence (suggested by the original correspondence between γ_1 and C_1) between two arcs of C_1 and corresponding arcs of γ; for w on each such arc of γ the conclusion of Lemma 4.5 as just interpreted is valid; the Lipschitz condition with respect to w' on the integrand of (4.20) holds uniformly for w on each of these two arcs of γ except for w in the neighborhoods of one end-point of each arc; these two neighborhoods correspond to one-sided neighborhoods of a single point on C_1, a point different from the exceptional point of C_1 under the original correspondence between γ_1 and C_1. Thus the inequalities involving τ_n are established uniformly for all w on γ_1 or γ_j, thus for all z on C_1 or C_j.

By taking exponentials, we now obtain

$$e^{-M} \leqq \frac{e^{ng}}{|\tau_n|} \leqq e^M, \qquad e^{-M} \leqq \left| \frac{e^{ng}(w^n - 1)}{\omega_{n-1}(z)} \right| \leqq e^M.$$

Another form of these inequalities is slightly more convenient:

$$\begin{aligned} &|\omega_{n-1}(z)| \leqq e^{M+ng} |w^n - 1| \leqq M_1 e^{ng}, \\ &|\omega_{n-1}(z)| \geqq M_2 e^{ng} |w^n - 1|, \end{aligned} \tag{4.21}$$

where $|w| = 1$ and z lies on C.[15] We have proved

LEMMA 4.6. *If C is a contour, inequalities* (4.21) *are uniformly valid, where* $|w| = 1$ *and z lies on C.*

For z on C_ρ, equation (4.1) may be written

$$\int_C \log |z - \zeta| \, du = \log \rho + g,$$

whence from Lemma 4.5

$$\left| \log \rho + g - \frac{1}{n} \log |\omega_{n-1}(z)| \right| \leqq \frac{M'}{n},$$

$$0 < M_1' e^{ng} \rho^n \leqq |\omega_{n-1}(z)| \leqq M_2' e^{ng} \rho^n, \qquad z \text{ on } C_\rho. \tag{4.22}$$

The derivation of (4.22) requires much less than that C be a contour, merely that the function

$$\log |z - \zeta| = \log |z - \psi(w)|$$

satisfy uniformly for all z on C_ρ a Lipschitz condition with respect to u or w, on each γ_j. It is sufficient (compare Lemma 4.9 below) if $|z - \psi(w)|$ satisfies such a Lipschitz condition, for which it is sufficient if $\psi(w)$ itself satisfies such a condition; the existence and continuity of $\psi'(w)$ on $|w| = 1$ are ample to ensure this; compare inequality (4.29) below.

Merely a part of the conclusions in (4.21) and (4.22) is summarized in

[15] Similar inequalities have been obtained by Curtiss [1] in the case that C is a single Jordan curve and where the mapping function $\psi(w)$ has a second derivative satisfying a Lipschitz condition of order $\alpha > 0$.

THEOREM 4.7. *Let C be a contour, and let the points z_k be equally distributed on C. Then we have*

$$\left|\frac{\omega_n(z)}{\omega_n(t)}\right| \leqq \frac{M_0}{\rho^n}, \qquad z \text{ on } C,\ t \text{ on } C_\rho,$$

where M_0 is a constant independent of n, z, and t.

Theorem 4.7 is one of the principal results of the present paper, to be applied several times below.

Since we have made Definition 4.2 fundamental in our present discussion, it is desirable to go into further detail concerning the properties of contours. Our next aim is to establish

THEOREM 4.8. *Let the function $|[\psi(w) - \psi(w')]/(w - w')|$ be uniformly bounded from zero for $|w| = 1$ and $|w'| = 1$, and on each closed arc of γ interior to a γ_j let*[16] *$\psi''(w)$ be continuous. Then C is a contour.*

In Theorem 4.8 (as in Theorem 4.10 below) we naturally assume an interpretation similar to that of Definition 4.2 relative to the various choices of the maps of C onto γ.

In the proof of Theorem 4.8 it will be convenient to have for reference

LEMMA 4.9. *If on a closed finite interval of the axis of reals or a closed arc of the unit circle the function $f(x_1, x_2)$ is positive and satisfies a Lipschitz condition of given order α $(0 < \alpha \leqq 1)$ in x_1 uniformly with respect to x_2:*

$$|f(x_1, x_2) - f(x_1', x_2)| \leqq M\,|x_1 - x_1'|^\alpha,$$

then $\log f(x_1, x_2)$ *satisfies on that interval or arc a Lipschitz condition of order α in x_1, uniformly with respect to x_2.*

Since $f(x_1, x_2)$ is positive and continuous on the given point set, that function is uniformly bounded from zero there: $f(x_1, x_2) > m > 0$. We have the expansion

$$\begin{aligned} &\log f(x_1, x_2) - \log f(x_1', x_2) \\ &\quad = \frac{f(x_1, x_2) - f(x_1', x_2)}{f(x_1', x_2)}\left\{1 - \frac{1}{2}\left[\frac{f(x_1, x_2) - f(x_1', x_2)}{f(x_1', x_2)}\right] + \cdots\right\}, \end{aligned} \tag{4.23}$$

which is valid provided $|[f(x_1, x_2) - f(x_1', x_2)]/f(x_1', x_2)| < 1$; this inequality is satisfied uniformly for x_1' sufficiently near x_1. Furthermore, it follows from the boundedness of $f(x_1, x_2)$ that the Lipschitz condition on $\log f(x_1, x_2)$ need be established only for $x_1' - x_1$ sufficiently small; compare (4.17). For $|x_1' - x_1| < \delta$ we have from (4.23) and from the Lipschitz condition on $f(x_1, x_2)$

[16] We denote as usual the p-th derivative of $f(z)$ by $f^{(p)}(z)$, and also use the notation $f^{0}(z) \equiv f(z)$, $f^{(1)}(z) \equiv f'(z)$, $f^{(2)}(z) \equiv f''(z)$.

In Theorem 4.8 and below, the requirement of the continuity of $\psi''(w)$ can be somewhat lessened; it is sufficient if $\psi''(w)$ is integrable and bounded, by obvious modifications of the proof given.

$$|\log f(x_1, x_2) - \log f(x_1', x_2)| \leqq \frac{M}{m}|x_1 - x_1'|^\alpha \cdot \left[1 + \frac{M}{2m}|x_1 - x_1'|^\alpha + \cdots\right],$$

provided $\delta^\alpha < m/M$. We read off at once the Lipschitz condition on log $f(x_1, x_2)$, including uniformity with respect to x_2.

In proving Theorem 4.8 we shall use the mean value theorem of Darboux [1]: *Let $f(z)$ and $f'(z)$ be continuous along the line segment from z_0 to z_1; then we have*

$$f(z_1) - f(z_0) = \lambda(z_1 - z_0) \cdot f'[z_0 + \epsilon(z_1 - z_0)], \tag{4.24}$$

where $0 \leqq \epsilon \leqq 1$, $|\lambda| \leqq 1$. If $f''(z)$ is also continuous, replacement of $f(z)$ by $f(z) - (z - z_1)f'(z)$ yields

$$f(z_1) = f(z_0) + f'(z_0)(z_1 - z_0) + \lambda(z_1 - z_0)^2 f''[z_0 + \epsilon(z_1 - z_0)]. \tag{4.25}$$

A method of proof of (4.24), quite different from that of Darboux, is the following. By integration along the segment from z_0 to z_1 we may write

$$f(z_1) - f(z_0) = \int_{z_0}^{z_1} f'(z) \cdot dz,$$

$$|f(z_1) - f(z_0)| \leqq |z_1 - z_0| \cdot \max |f'(z)|; \tag{4.26}$$

this is essentially equation (4.24).[17] If the curve (assumed to have a continuously turning tangent) on which we study $f(z)$ is no longer the line segment, inequality (4.26) remains valid provided we replace $|z_1 - z_0|$ by the length of the curve, whence

$$f(z_1) - f(z_0) = \lambda f'(z') \int_{z_0}^{z_1} |dz|, \tag{4.27}$$

where z' is some point on the curve. Likewise equation (4.25) may be replaced by

$$f(z_1) = f(z_0) + f'(z_0)(z_1 - z_0) + \lambda f''(z')\left(\int_{z_0}^{z_1} |dz|\right)^2, \tag{4.28}$$

where we have $|\lambda| \leqq 1$, where the integral is taken over the curve considered, and where z' lies on this curve.

If the function $f(z)$ is studied *on a circular arc* on which z_0 and z_1 lie, we may use the inequality

$$\int_{z_0}^{z_1} |dz| \leqq \tfrac{1}{2}\pi |z_1 - z_0|;$$

we assume here that the arc considered is not greater than a semicircle; in the contrary case we construct the subsequent proof in several steps as in treating (4.17). Thus (4.27) takes the form

[17] It is possible to establish (4.24) where $0 < \epsilon < 1$ by reasoning slightly more refined than that given, but this distinction does not concern us in the present paper.

$$f(z_1) - f(z_0) = \lambda(\tfrac{1}{2}\pi)(z_1 - z_0)f'(z'), \tag{4.29}$$

and (4.28) takes the form

$$f(z_1) = f(z_0) + f'(z_0)(z_1 - z_0) + \lambda(\tfrac{1}{2}\pi)^2(z_1 - z_0)^2 f''(z'), \tag{4.30}$$

with $|\lambda| \leqq 1$.

Under the hypothesis of Theorem 4.8 we now write

$$\Psi(w, w') = \frac{\psi(w) - \psi(w')}{w - w'}, \tag{4.31}$$

$$\Psi'(w, w') = \frac{(w - w')\psi'(w) - [\psi(w) - \psi(w')]}{(w - w')^2}, \tag{4.32}$$

where the accent indicates differentiation with respect to w. Let the points w and w' both lie on a closed subarc interior to γ_j; then it follows from (4.30) and (4.32) that $\Psi'(w, w')$ is uniformly bounded unless $w = w'$. Considered as a function of w, the function $\Psi(w, w')$ is continuous, even in the point $w = w'$ provided we set $\Psi(w', w') = \psi'(w')$.

In the derivation of (4.27) and (4.28), the existence or non-existence of $f'(z)$ or $f''(z)$ at a single point may be disregarded, provided $f'(z)$ or $f''(z)$ exists otherwise, is uniformly bounded, and is continuous. It then follows from (4.29) that $\Psi(w, w')$ satisfies a Lipschitz condition

$$|\Psi(w_1, w') - \Psi(w_2, w')| \leqq L\,|w_1 - w_2|$$

for w_1, w_2, w' on the closed subarc interior to γ_j, or even if w' lies on a closed subarc interior to a γ_k, $k \neq j$. The Lipschitz condition in w on the function (4.10) uniformly with respect to w' now follows from Lemma 4.9 by means of the inequality

$$\big|\,|\Psi(w_1, w')| - |\Psi(w_2, w')|\,\big| \leqq |\Psi(w_1, w') - \Psi(w_2, w')|;$$

the fact that $\Psi(w, w')$ is bounded from zero is part of our hypothesis.

The existence of $\psi''(w)$ on $|w| = 1$ implies the continuity of $\psi'(w)$ there, hence implies that C_j has a continuously turning tangent. It follows from the method of proof of (4.13) that $\psi'(w)$ satisfies a Lipschitz condition on the arc considered. It remains, however, to prove boundedness of the function (4.10) for $|w| = 1$ and $|w'| \geqq 1$. We proceed as in the proof and application of Lemma 4.4; the inequality $\psi'(w) \neq 0$ is a consequence of the hypothesis that $\Psi(w, w')$ is bounded from zero. It now follows from the theorem of Kellogg [1] that $\psi'(w)$ exists and is continuous in the two-dimensional sense in the neighborhood of $|w| = 1$. Since $\psi'(w)$ is different from zero on $|w| = 1$, it follows that $\psi'(w)$ is different from zero in the neighborhood of $|w| = 1$. Use of the transformations $w_1 = 1/w$, $w_1' = 1/w'$ and then equation (4.24) for the segment w_1w_1' shows that the function

$$\frac{\psi(w) - \psi(w')}{w - w'}$$

is uniformly bounded for $|w| = 1$ and $|w'| \geqq 1$ in the neighborhood of points of continuity of $\psi(w)$; we use here the boundedness of $\psi'(w)$ and the fact that in the derivation of (4.24) the boundedness and continuity of the derivative are sufficient even if the derivative fails to exist at a single point.[18] Further use of the transformations $w_1 = 1/w$, $w_1' = 1/w'$ gives a convenient path, namely, the line segment joining w_1 and w_1', for use; the transformations $z = \psi(w)$, $z' = \psi(w')$ then show by means of (4.27) that the function

$$\frac{w - w'}{\psi(w) - \psi(w')} = \frac{\varphi(z) - \varphi(z')}{z - z'}$$

is uniformly bounded at points of continuity of $\psi(w)$, for $|w| = 1$ and $|w'| \geqq 1$. Thus the function (4.10) is locally bounded for $|w| = 1$ and $|w'| \geqq 1$ in the neighborhood of points of continuity of $\psi(w)$. Theorem 4.8 is established.

Still another sufficient condition may be mentioned:

THEOREM 4.10. *If the function*

$$\log\left[\frac{\psi(w) - \psi(w')}{w - w'}\right] \tag{4.33}$$

satisfies on each closed arc interior to every γ_j a Lipschitz condition, then C is a contour.

The Lipschitz condition on the function (4.33) implies at once the corresponding Lipschitz condition on the function (4.10), the real part of (4.33). By the method of proof of (4.13) and by Lemma 4.3 it follows that $\psi'(w)$ exists on any closed arc interior to γ_j and satisfies a Lipschitz condition there, and moreover that $\psi'(w)$ is different from zero. The boundedness of the function (4.10) follows from Kellogg's theorem by means of (4.17a) and the later reasoning already given. This completes the proof of Theorem 4.10.

5. **Tchebycheff, Faber, and orthogonal polynomials.** Let C be an arbitrary Jordan curve. Let $T_n(z) \equiv z^n + a_1 z^{n-1} + \cdots + a_n$ be the uniquely determined polynomial of that form for which

$$\max\,[\,|T_n(z)|,\ z \text{ on } C]$$

[18] This method enables us to establish the following theorem:

Let C be an analytic Jordan arc (this restriction can be lightened); *let the function $\psi(z)$ be analytic in a one-sided neighborhood of C, continuous in the corresponding closed region R. For a particular interior point z' of C let $\psi'(z')$ exist on C in the one-dimensional sense, and suppose*

$$\limsup_{z \to z',\ z \text{ interior to } R} \psi'(z)$$

is bounded. Then the difference-quotient $[\psi(z) - \psi(z')]/(z - z')$ is bounded in the two-dimensional sense for z interior to R, and the derivative $\psi'(z')$ is a two-dimensional derivative.

The last part of this theorem is a consequence of a theorem of Lindelöf.

is least; then $T_n(z)$ is called the *Tchebycheff polynomial of degree n belonging to C*; of course the corresponding terminology is appropriate also if C is replaced by its closed interior $\bar{C}$; the Tchebycheff polynomials of degree n belonging to C and $\bar{C}$ are identical.

Faber [2] has established the essence of

THEOREM 5.1. *Let C be an analytic Jordan curve, and let $T_n(z)$ be the Tchebycheff polynomial of degree n belonging to C; then we have*

$$\left|\frac{T_n(z)}{T_n(t)}\right| \leqq \frac{M}{\rho^n}, \qquad z \text{ on } C,\ t \text{ on } C_\rho,$$

where M is a constant independent of n, z, and t.

Here and henceforth (Theorems 5.3, 5.5) we assume ρ so chosen that no zero of one of the polynomials involved lies on C_ρ.

Theorem 5.1 is an immediate consequence of Faber's equation involving the usual mapping function $\varphi(z)$:

$$\lim_{n\to\infty} \frac{T_n(z)}{e^{ng}[\varphi(z)]^n} = 1, \tag{5.1}$$

uniformly for z on or exterior to C.

Let now E be chosen as the closed segment $-1 \leqq z \leqq 1$ of the axis of reals. The Tchebycheff polynomial (defined as above) belonging to E of degree n is

$$T_n(z) \equiv \frac{1}{2^{n-1}} \cos\,(n \cos^{-1} z);$$

indeed, Faber [2] shows that the Faber and Tchebycheff polynomials (the Faber polynomials are defined below) both for this segment and for all ellipses with the foci $+1$ and -1 are given by this formula. The zeros of $T_n(z)$ are known to lie on E (see for instance Pólya-Szegö [1], vol. II, pp. 75 and 266). It follows from this formula and from Faber's relation (5.1) for z exterior to E that we have uniformly on any closed set exterior to E

$$\lim_{n\to\infty} \frac{2^n T_n(z)}{[\varphi(z)]^n} = 1,$$

since $e^g = \frac{1}{2}$ for the segment. We have established the well-known

THEOREM 5.2. *If E is the segment $-1 \leqq z \leqq 1$, the Tchebycheff polynomial $T_n(z)$ of degree n belonging to E satisfies the inequality*

$$\left|\frac{T_n(z)}{T_n(t)}\right| \leqq \frac{M}{\rho^n}, \qquad -1 \leqq z \leqq 1,\ t \text{ on } C_\rho,$$

where M is a constant independent of n, t, and z.

For an analytic Jordan curve Faber [1] has studied also other polynomials $F_n(z)$ of respective degree n, which we shall call the Faber polynomials. He establishes the relation

$$\lim_{n\to\infty} \frac{F_n(z)}{[\varphi(z)]^n} = 1 \tag{5.2}$$

uniformly for z on or exterior to C. We shall proceed to derive a slightly weaker relation, valid even if C is not analytic.

DEFINITION 5.1. If C is a Jordan curve, and if in the notation already introduced the function

$$\frac{w\psi'(w)}{\psi(w) - \psi(w')} - \frac{w}{w - w'} \tag{5.3}$$

is bounded on $|w| = 1$ uniformly with respect to w' for $|w'| = 1$, and if (5.3) is expressed for $|w| > 1$ and $|w'| = 1$ by the Cauchy integral over $|w| = 1$, we shall say that C is *of type A*.

Faber's polynomials $F_n(z)$ are defined by the identity

$$\frac{w\psi'(w)}{\psi(w) - z} = 1 - \sum_{\mu=0}^{\infty} F_{\mu+1}(z)\cdot w^{-\mu-1}, \tag{5.4}$$

an expansion valid for $|w| > |\varphi(z)|$, thus valid for $|w| = \rho > 1$ and for z on C. A comparison function can be written

$$\frac{w}{w - w'} = \sum_{\mu=0}^{\infty}\left(\frac{w'}{w}\right)^{\mu}, \qquad |w| > |w'|.$$

The coefficient of $w^{-\mu-1}$ is bounded here for all w' even if we take $|w'| = |w| = 1$.

If the conditions of Definition 5.1 are fulfilled, it follows from the Cauchy inequality that the coefficients in the expansion of (5.3) in powers of w are bounded uniformly with respect to w' for all $|w'| = 1$, so the functions $F_n(z)$ defined by (5.4) are also bounded uniformly for all $|w'| = 1$, where $z = \psi(w')$, that is to say, are bounded uniformly for z on C. The uniform boundedness on C_ρ of $\rho^n/F_n(z)$ follows as proved by Faber, so we have established

THEOREM 5.3. *If C is a curve of type A, we have for the Faber polynomials $F_n(z)$:*

$$\left|\frac{F_n(z)}{F_n(t)}\right| \leq \frac{M}{\rho^n}, \qquad z \text{ on } C,\ t \text{ on } C_\rho,$$

where M is a constant independent of n, t, and z.

In relation to Theorem 5.3 it is of interest to prove

THEOREM 5.4. *Let C be a Jordan curve, let $\psi'(w)$ be different from zero on $|w| = 1$, and let $\psi''(w)$ exist and be continuous on $|w| = 1$. Then C is of type A.*

We write (5.3) in the form

$$\frac{-w[\psi(w) - \psi(w') - (w - w')\psi'(w)]}{(w - w')[\psi(w) - \psi(w')]},$$

which by (4.30) can be expressed for $|w| = 1$ and $|w'| = 1$

$$\lambda\left(\frac{\pi}{2}\right)^2 \frac{w(w - w')\psi''(w_0)}{\psi(w) - \psi(w')}, \qquad |\lambda| < 1.$$

The continuity of $\psi'(w)$ follows from the existence of $\psi''(w)$; use of the transformation $w = \varphi(z)$ and equation (4.27) implies that $(w - w')/[\psi(w) - \psi(w')]$ is uniformly bounded on C; consequently the function (5.3) is uniformly bounded for w and w' on $|w| = 1$ and $|w'| = 1$. As in the proof of Theorem 4.8 it now follows that $\psi'(w)$ is continuous and different from zero throughout the neighborhood of $|w| = 1$, hence throughout the closed region $|w| \geqq 1$. The continuity of (5.3) in the two-dimensional sense considered as a function of w for $|w| = 1$ and for w' fixed with $|w'| = 1$, $w \neq w'$, follows at once.

With the factor w suppressed, the function (5.3) may be written in the form

$$\frac{f_{w'}(w) - f_{w'}(w')}{w - w'}, \tag{5.3'}$$

$$f_{w'}(w) \equiv \frac{w - w'}{\psi(w) - \psi(w')}\,\psi'(w), \qquad f_{w'}(w') \equiv 1.$$

Continuity of $\psi'(w)$ in the two-dimensional sense and use of (4.26) after the transformation $w_1 = 1/w$ together with Lindelöf's theorem show that $\psi'(w)$ on $|w| = 1$ is a two-dimensional derivative; compare the latter part of the proof of Theorem 4.8. The function (5.3′) is uniformly bounded for w and w' on $|w| = 1$ and $|w'| = 1$; the function $f_{w'}(w)$ considered as a function of w for w' fixed on $|w'| = 1$ is continuous for $|w| \geqq 1$. It follows from Theorem 10.1 below that the function (5.3′) is uniformly bounded for all $|w'| = 1$ and for all $|w| \geqq 1$. This is sufficient for the expression of (5.3) as the Cauchy integral over $|w| = 1$, valid for $|w| > 1$ and $|w'| = 1$. Theorem 5.4 is established.

Let now C be an analytic Jordan curve, and let $w(z)$ be positive and continuous on C. Szegö ([1]; §§16.4, 16.5) defines polynomials $Q_n(z)$ of respective degrees $n = 0, 1, 2, \cdots$, such that the coefficient of z in $Q_n(z)$ is real and positive, polynomials normal and orthogonal on C in the sense that the integral

$$\frac{1}{L}\int_C Q_m(z)\overline{Q_n(z)}w(z)\,|dz|, \qquad L = \int_C |dz|,$$

is equal to unity or zero according as m and n are or are not equal. Szegö then proves that $Q_n(z)$ is given asymptotically by

$$Q_n(z) \sim \left(\frac{L}{2\pi}\right)^{\frac{1}{2}} [\varphi'(z)]^{\frac{1}{2}}[\varphi(z)]^n/\Delta(z), \tag{5.5}$$

where $w = \varphi(z)$ is the function previously considered, and where $\Delta(z)$ is analytic and different from zero in K even at infinity, positive at infinity, and where $|\Delta(z)|^2$ is continuous in $K + C$ and equal to $w(z)$ on C. The asymptotic rela-

tion (5.5) is valid uniformly for z on any closed set exterior to C, in the sense that the ratio of the two members of (5.5) approaches unity uniformly. If $\Delta(z)$ is analytic in the closed exterior of C, the relation (5.5) is valid uniformly also on C and even in a sufficiently small neighborhood in the interior of C.

As a first consequence of (5.5) it may be remarked that although the roots of $Q_n(z)$ need not lie in E, those roots must lie for n sufficiently large interior to any preassigned C_ρ. A second consequence of (5.5) is formulated as

THEOREM 5.5. *If C is an analytic Jordan curve, we have the uniform inequality*

$$0 < m\rho^n < |Q_n(z)|, \qquad z \text{ on } C_\rho; \tag{5.6}$$

if also $\Delta(z)$ is analytic in the closed exterior of C, we have the uniform inequality

$$\left|\frac{Q_n(z)}{Q_n(t)}\right| \leqq \frac{M}{\rho^n}, \qquad z \text{ on } C,\ t \text{ on } C_\rho. \tag{5.7}$$

Inequalities (5.6) and (5.7) enable us to apply Theorems 3.4 and 3.2 respectively.

If now E is taken as the closed segment $-1 \leqq z \leqq 1$ of the axis of reals, with weight function unity, the orthogonal polynomials are those of Legendre:

$$\int_{-1}^{1} P_m(z)P_n(z)\,dz = 2/(2n+1) \quad \text{or } 0$$

according as m and n are or are not equal. The asymptotic formula is well known and may be taken as

$$P_n(z) \sim (2\pi n)^{-\frac{1}{2}}(z^2-1)^{-\frac{1}{4}}[z+(z^2-1)^{\frac{1}{2}}]^{n+\frac{1}{2}},$$

valid uniformly on any closed set having no point in common with E. On E itself there is the inequality $|P_n(z)| \leqq 1$.

THEOREM 5.6. *If E is the segment $-1 \leqq z \leqq 1$, we have the uniform inequality for the orthogonal polynomials $P_n(z)$:*

$$0 < m\rho^n n^{-\frac{1}{2}} < |P_n(z)|, \qquad z \text{ on } C_\rho. \tag{5.8}$$

Moreover, we have the uniform inequality

$$\left|\frac{P_n(z)}{P_n(t)}\right| \leqq \frac{Mn^{\frac{1}{2}}}{\rho^n}, \qquad z \text{ on } E,\ t \text{ on } C_\rho.$$

We remark too that a result similar to (5.8) holds for the polynomials $p_n(z)$ normal and orthogonal on E: $-1 \leqq z \leqq 1$ with respect to a weight function, which is arbitrary except for certain continuity conditions (Szegö [1], p. 290):

$$0 < m\rho^n < |p_n(z)|, \qquad z \text{ on } C_\rho. \tag{5.9}$$

Inequalities (5.8) and (5.9) are sufficient for the application of Theorem 3.4.

6. **Poles of lemniscates.** A *lemniscate* Γ_μ is defined as a locus of a point z under the conditions

$$(6.1)\qquad \Gamma_\mu : |\, p(z)\, | = \mu > 0, \qquad p(z) \equiv (z - \beta_1)(z - \beta_2) \cdots (z - \beta_\lambda),$$

the numbers μ and β_k being fixed. A lemniscate then consists of a finite number of Jordan curves which are mutually exterior except possibly for a finite number of points each of which may belong to several of the Jordan curves; the Jordan curves are analytic except at such common points; see for instance Walsh [1].

It is natural to study a polynomial of degree $n = m\lambda - 1$ interpolating to a given function in each of the points β_k counted m times. Then we have in the usual notation for $\omega_n(z) \equiv [p(z)]^m$

$$(6.2)\qquad \left|\frac{\omega_n(z)}{\omega_n(t)}\right| = \left(\frac{\mu}{\mu_1}\right)^m, \qquad z \text{ on } \Gamma_\mu,\ t \text{ on } \Gamma_{\mu_1}.$$

The mapping function $w = \varphi(z)$ for the exterior of the lemniscate Γ_μ is precisely $[p(z)/\mu]^{1/\lambda}$, so C_ρ is the locus $|\, p(t)\, | = \mu\rho^\lambda$, and (6.2) can be written

$$(6.3)\qquad \left|\frac{\omega_n(z)}{\omega_n(t)}\right| = \frac{1}{\rho^{m\lambda}} = \frac{1}{\rho^{n+1}}, \qquad z \text{ on } C,\ t \text{ on } C_\rho.$$

If we relinquish the requirement $n = m\lambda - 1$, we set $n + 1 = q\lambda + r$ $(0 \leqq r < \lambda)$; the points of interpolation shall be the β_k each counted q times, and in addition the points $\beta_1, \beta_2, \cdots, \beta_r$. We may write for z on Γ_μ and t on Γ_{μ_1}

$$\left|\frac{\omega_n(z)}{\omega_n(t)}\right| = \left(\frac{\mu}{\mu_1}\right)^q \cdot \left|\frac{(z - \beta_1) \cdots (z - \beta_r)}{(t - \beta_1) \cdots (t - \beta_r)}\right|.$$

The second factor in the right member is uniformly bounded, so we have

$$(6.4)\qquad \left|\frac{\omega_n(z)}{\omega_n(t)}\right| \leqq M\left(\frac{\mu}{\mu_1}\right)^q \leqq M_1\left(\frac{\mu}{\mu_1}\right)^{n/\lambda}, \qquad z \text{ on } \Gamma_\mu,\ t \text{ on } \Gamma_{\mu_1}.$$

If C denotes Γ_μ and C_ρ denotes Γ_{μ_1}, this last member may be written M_1/ρ^n.

Chapter III. Degree of convergence, Problem β

7. **Jordan curves.** We now reach the principal results of the present paper, notably application of the formulas and inequalities of §3 combined with the inequalities on polynomials considered in §§4–6, to derive results on degree of convergence on a set E when the given function is known to be analytic not merely on E but also throughout the interior of some C_ρ, and is assumed to have certain continuity properties on C_ρ. We shall have frequent occasion to employ the following theorem due to Curtiss [2] and relating to Problem α:[19]

THEOREM 7.1. *Let E be a closed limited point set bounded by a finite number of mutually exterior analytic Jordan curves C. Let $f(z)$ be continuous on E, analytic in the interior points of E, and let $f^{(p)}(z)$ for $p \geqq 0$ defined in the one-dimensional sense on C be continuous on C and satisfy on C a Lipschitz condition*

[19] If C consists of a single analytic Jordan curve, this conclusion (7.1) can be established for the case $p = 0$, $0 < \alpha < 1$, by choosing the $p_n(z)$ as the arithmetic mean of order n corresponding to the development of $f(z)$ in the Faber polynomials belonging to C.

of order α $(0 < \alpha \leqq 1)$. *Then there exist polynomials* $p_n(z)$ *of respective degrees* $n = 1, 2, \cdots$, *such that*

$$(7.1) \qquad |f(z) - p_n(z)| \leqq \frac{M}{n^{p+\alpha}}, \qquad z \text{ on } E,$$

where M *is a constant independent of* n *and* z.[20]

It is of interest in connection with Theorem 7.1 to study the relation between one-dimensional derivatives and two-dimensional derivatives on C, and also to study the continuity in the two-dimensional sense of such derivatives; this study we postpone until §10. We remark here, however, that in proving Theorem 7.1 Curtiss assumes that $f^{(p)}(z)$ is continuous on C in the two-dimensional sense, which we prove in §10 to be a consequence of the continuity of $f(z)$ on E and the continuity of $f^{(p)}(z)$ on C in the one-dimensional sense.

Let C be a Jordan curve or several Jordan curves whose interiors are mutually disjoint. We say that $f(z)$ is *of class* $L(p, \alpha)$ *on* C, where p is a non-negative integer and where $0 < \alpha \leqq 1$, provided $f(z)$ is analytic interior to each of the Jordan curves composing C, is continuous on the corresponding closed set, and if the function $f^{(p)}(z)$ exists and satisfies on C a Lipschitz condition of given order α.

An immediate consequence of Theorems 3.2, 4.7, and 7.1 is

THEOREM 7.2. *Let* E *be a closed limited point set bounded by a contour* C, *and let* C_ρ *also be a contour.*[21] *Let the function* $f(z)$ *be of class* $L(p, \alpha)$ $(0 < \alpha \leqq 1)$ *on* C_ρ. *Let* $p_n(z)$ *denote the unique polynomial of degree* n *which interpolates to* $f(z)$ *in* $n + 1$ *equally distributed points of* C. *Then we have*

$$(7.2) \qquad |f(z) - p_n(z)| \leqq \frac{M}{n^{p+\alpha}\rho^n}, \qquad z \text{ on } E,$$

where M *is a constant independent of* n *and* z.

[20] We mention now an incidental application of Theorem 7.1 to Problem α. Jackson [2] under the present hypothesis on C and E assumes that $\rho(z)$ is a non-negative measurable bounded function with a positive lower bound. If $p > 0$ is given, he shows that if $f(z)$ is continuous on E, analytic in the interior points of E, and if there exist polynomials $p_n(z)$ of respective degrees n such that $|f(z) - p_n(z)| \leqq \epsilon_n$ on E, then the polynomials $L_n(z)$ of respective degrees n which minimize

$$\int_C \rho(z)|f(z) - L_n(z)|^p |dz|$$

satisfy the inequality $|f(z) - L_n(z)| \leqq Mn^{1/p}\epsilon_n$, z on E. If $f(z)$ satisfies the conditions of Theorem 7.1, we therefore have $|f(z) - L_n(z)| \leqq M/n^{p+\alpha-1/p}$, z on E.

A similar remark applies to the corollary to Theorem 12.1.

[21] Throughout the present paper we suppose C_ρ to have no multiple points, for Theorem 7.1 has never been established for the case of Jordan curves with multiple points.

But results in the direction of a converse of Theorem 7.1 (such as Theorem 12.2 below) present no great difficulty even if C_ρ has multiple points; indeed, the original method of proof shows that *on an analytic Jordan arc* a suitable degree of convergence by polynomials implies condition (12.14) on any closed subarc.

In §5 we indicated that for a suitably restricted Jordan curve C some roots of the Tchebycheff, Faber, and orthogonal polynomials of degree n belonging to C may, for small n, lie on or exterior to a given C_ρ; but when C_ρ is given, all roots of such polynomials for n sufficiently large lie interior to C_ρ. For the validity of the usual proof of (3.3), however, it is necessary that *all* the points of interpolation lie interior to C_ρ. Consequently, if $f(z)$ is given analytic interior to a particular C_ρ, continuous on the corresponding closed set, the polynomials $p_n(z)$ of interpolation in the roots of the Tchebycheff, Faber, and orthogonal polynomials need not be defined when n is small, but surely are defined when n is sufficiently large. This restriction on n is ordinarily not serious, however, for we may define $f(z)$ artificially in any way at the points of interpolation on or exterior to C_ρ; such an inequality as (7.2) or (7.3) below, if valid for n sufficiently large, remains valid for all n provided the constant M is suitably altered. This convention is to be understood in the immediately succeeding theorems.

THEOREM 7.3. *Let E be the closed interior of a Jordan curve C, let the function $f(z)$ be of class $L(p, \alpha)$ on C_ρ $(0 < \alpha \leqq 1)$. Let C be analytic and consider the Tchebycheff polynomials $T_n(z)$ belonging to C, or let C be of type A and consider the Faber polynomials $F_n(z)$ belonging to C, or let C be analytic and consider the polynomials $Q_n(z)$ normal and orthogonal on C with respect to a positive analytic weight function. In these various cases let $p_n(z)$ denote the polynomial of degree n which interpolates to $f(z)$ in the $n + 1$ roots of $T_{n+1}(z)$, $F_{n+1}(z)$, or $Q_{n+1}(z)$. Then we have for n sufficiently large*

$$(7.3) \qquad |f(z) - p_n(z)| \leqq \frac{M}{n^{p+\alpha}\rho^n}, \qquad z \text{ on } E,$$

where M is independent of n and z.

Theorem 7.3 is a consequence of Theorems 3.2 and 7.1, taken in the various cases together with Theorems 5.1, 5.3, and 5.5. In the last case, interpolation in the roots of the $Q_{n+1}(z)$, where we still suppose the weight function positive and analytic on the analytic Jordan curve C, even more can be said of the convergence of the sequence $p_n(z)$. Contrary to our usual custom, let us for the moment use the notation C_ν for the locus $|\varphi(z)| = \nu$ even if ν is less than unity; the function $\varphi(z)$ is analytic on C, so C_ν now has a meaning for $\nu < 1$ but $1 - \nu$ sufficiently small. It follows from the fact that the asymptotic formula (5.6) is valid even in a sufficiently small neighborhood interior to C, that we have also for n sufficiently large and for $\nu < 1$ but $1 - \nu$ sufficiently small

$$(7.4) \qquad |f(z) - p_n(z)| \leqq \frac{M_1\nu^n}{n^{p+\alpha}\rho^n}, \qquad z \text{ on } C_\nu,$$

where M_1 is independent of n and z. The methods already used yield (7.4) as well as (7.3).

Another result on interpolation in the roots of the $Q_n(z)$ follows at once from Theorems 3.4, 5.5, and 7.1:

THEOREM 7.4. *Let E be the closed interior of an analytic Jordan curve C, let the function $f(z)$ be of class $L(p, \alpha)$ $(0 < \alpha \leqq 1)$ on C_ρ. Let $p_n(z)$ denote the polynomial of degree n which interpolates to $f(z)$ in the $n + 1$ roots of the polynomial $Q_{n+1}(z)$ of the set normal and orthogonal on C with respect to a positive and continuous weight function. Then we have for n sufficiently large*

$$\int_C |f(z) - p_n(z)|^2 \cdot |dz| \leqq \frac{M}{n^{2p+2\alpha}\rho^{2n}},$$

when M is independent of n and z.

Of course the boundedness of

$$\int_C |Q_n(z)|^2 \cdot |dz|,$$

used in the proof of Theorem 7.4, follows from the boundedness on C of the reciprocal of the weight function and the normality on C of the $Q_n(z)$.

We study the case of orthogonal polynomials in more detail, using now for simplicity the norm function unity. Let us say that a function $f(z)$ analytic interior to a rectifiable Jordan curve C and continuous[22] in the corresponding closed region is of class F_q on C if the numbers

$$n^q \cdot \int_C f(z)\overline{Q_n(z)}\,|dz| \qquad (n = 0, 1, 2, \cdots; q > 0)$$

are uniformly bounded. If we set

$$a_k = \int_C f(z)\overline{Q_k(z)}\,|dz|, \qquad S_n(z) = \sum_{k=0}^{n} a_k Q_k(z),$$

it follows (Walsh [1], pp. 108, 112, 119) that $S_n(z)$ converges in the mean to $f(z)$ on C and that we have

$$\int_C |f(z) - S_n(z)|^2 |dz| = \sum_{n+1}^{\infty} |a_k|^2.$$

Consequently, if $f(z)$ is of class F_q on C, we have for $q > \frac{1}{2}$

$$\int_C |f(z) - S_n(z)|^2 |dz| \leqq \frac{M}{n^{2q-1}}, \tag{7.5}$$

a result on Problem α.[23]

This definition of class F_q is similar to the notion of *order* of a function $f(z)$ defined by Hadamard [1] and employed also by Mandelbrojt ([1], Chapter VII; [2]) when C is the unit circle. If $f(z) = \sum a_k z^k$, $|z| < 1$, the order ω of $f(z)$ on C is defined as

[22] In certain cases this requirement of continuity may be somewhat lessened, provided boundary values exist in a certain sense. Compare Smirnoff [1], Keldysch and Lavrentieff [1], as well as the study of class H_2 by Walsh [1].

[23] It is obviously possible to study also orthogonality as measured by a surface integral and to obtain similar results.

$$\omega = \limsup_{k\to\infty} \frac{\log |a_k|}{\log k} + 1,$$

so $\omega - 1$ and q are closely related.

Inequality (7.5) taken together with Theorems 3.3 and 4.7 gives

THEOREM 7.5. *Let E be the closed interior of a Jordan curve C which is a contour. Let $f(z)$ be analytic interior to C_ρ, continuous in the corresponding closed region, and of class F_q on C_ρ. Let $p_n(z)$ be a polynomial of degree n interpolating to $f(z)$ in $n + 1$ equally distributed points on C. Then we have*

$$|f(z) - p_n(z)| \leqq \frac{M}{n^{q-\frac{1}{2}}\rho^n}, \qquad z \text{ on } C, \tag{7.6}$$

where M is independent of n and z.

The results analogous to Theorem 7.5 are valid and immediate, when under suitable hypothesis on C we interpolate to $f(z)$ in the roots of the Tchebycheff, Faber, or orthogonal polynomials belonging to C.[24]

It happens that we establish the same inequalities (7.2) and (7.6) for a function analytic interior to C, continuous in the corresponding closed region, which satisfies a Lipschitz condition of order $q - \frac{1}{2}$ on C as for a similar function of class F_q, assuming $0 < q - \frac{1}{2} \leqq 1$. It is not to be supposed, however, that these two classes are identical.

Thus, let $f(z) \equiv (1 + z)^\alpha$ $(0 < \alpha < \frac{1}{2})$, where we choose the branch which is real when z is real and positive; then $f(z)$ is analytic for $|z| < 1$, continuous for $|z| \leqq 1$. The function $f(z)$ satisfies on $C: |z| = 1$ a Lipschitz condition of no order $\beta > \alpha$, and yet the binomial theorem[25] shows $f(z)$ to be of class $F_{1+\alpha}$ on $|z| = 1$.

However, it is true that if $f(z)$ is analytic for $|z| < 1$, continuous for $|z| \leqq 1$, and satisfies on $C: |z| = 1$ a Lipschitz condition of order α $(0 < \alpha \leqq 1)$, then $f(z)$ is of class F_α; this follows from the analogous theorem concerning Fourier coefficients, which is well known; also if $f^{(p)}(z)$ satisfies a Lipschitz condition of order α on C $(0 < \alpha \leqq 1)$, then $f(z)$ is of class $F_{p+\alpha}$. Likewise if C is an analytic Jordan curve and $f(z)$ is of class $L(p, \alpha)$ on C it is true that $f(z)$ is of class $F_{p+\alpha}$. For we have from Theorem 7.1 in the notation already introduced

$$\int_C |f(z) - S_n(z)|^2 |dz| = \sum_{k=n+1}^{\infty} |a_k|^2 \leqq \int_C |f(z) - p_n(z)|^2 |dz| \leqq \frac{M_1}{n^{2p+2\alpha}},$$

whence $|a_n| \leqq M/n^{p+\alpha}$.

[24] We remark that for the case that C_ρ is the unit circle and $p_n(z)$ the polynomial interpolating to $f(z)$ in the origin (which is also the n-fold root of the Tchebycheff, Faber, and orthogonal polynomial with unit weight function), inequality (7.6) can be improved. Suppose $f(z) \equiv \sum_{k=0}^{\infty} a_k z^k$, where $|a_k| \leqq M/k^q$; then on $C: |z| = 1/\rho$ we have

$$|f(z) - p_n(z)| = \Big|\sum_{k=n+1}^{\infty} a_k z^k\Big| \leqq \frac{M_1}{n^q} \sum_{k=n+1}^{\infty} |z| \leqq \frac{M_0}{n^q \rho^n}.$$

[25] See Dienes [1], p. 24.

The significance of the results established in §7 is two-fold, as has already been remarked: (i) we have determined inequalities for the convergence of specific sequences of polynomials, (ii) we have existence theorems for sequences converging with at least a certain degree of convergence.

In the special case that the curve C of Theorem 7.2 is the unit circle, the same degree of convergence as is indicated by (7.2) is exhibited by the partial sums of the Taylor development of the function $f(z)$ (see Walsh and Sewell [3]). This is a further result concerning item (i), and likewise a method different from those presented in Theorems 7.2–7.5 for treating item (ii). It may be mentioned, however, that for the unit circle C interpolation in the roots of Tchebycheff, Faber, or orthogonal polynomials (with norm function unity) is also interpolation in the origin, and leads to the Taylor development; in this sense, the further result just mentioned is contained in Theorem 7.3. Still another method for establishing this degree of convergence (7.2) as proved in Theorem 7.2 for C the unit circle is the direct comparison (Walsh [1], §7.1) of degrees of convergence of polynomials interpolating in the origin and polynomials interpolating in the roots of unity.

A simple conformal transformation enables us to use item (i) and consequently item (ii) for the unit circle in studying corresponding results for the segment $-1 \leqq z \leqq 1$. This method has previously been used (compare Walsh and Sewell [3]) and satisfactory results obtained. Hence we shall not carry out this method in the present paper, but shall continue with the application of the methods of §3. In §8, then, the emphasis is on item (i) rather than item (ii).

8. The segment $-1 \leqq z \leqq 1$.

The special interest of the results about to be proved is that they complement the classical results[26] on real variables.

Theorems 3.2, 5.2, and 7.1 yield

THEOREM 8.1. *Let E be the segment $-1 \leqq z \leqq 1$, so that C_ρ is the ellipse with foci -1 and $+1$ and with semi-sum of axes $\rho > 1$. Let $f(z)$ be of class $L(p, \alpha)$ $(0 < \alpha \leqq 1)$ on C_ρ. Let $p_n(z)$ be the polynomial of degree n which interpolates to $f(z)$ in the $n + 1$ roots of the Tchebycheff polynomial* $\cos[(n + 1)\cos^{-1} z]/2^n$. *Then we have*

$$|f(z) - p_n(z)| \leqq \frac{M}{n^{p+\alpha}\rho^n} \qquad (-1 \leqq z \leqq 1),$$

where M is a constant independent of n and z.

Theorems 3.1, 5.6, and 7.1 yield

THEOREM 8.2. *Let E and $f(z)$ satisfy the hypothesis of Theorem* 8.1 *and let $p_n(z)$ be the polynomial of degree n which interpolates to $f(z)$ in the $n + 1$ roots of the Legendre polynomial $P_{n+1}(z)$. Then we have*

[26] Compare Bernstein [2] and de la Vallée Poussin [1].

$$|f(z) - p_n(z)| \leqq \frac{M}{n^{p+\alpha-\frac{1}{2}}\rho^n} \qquad (-1 \leqq z \leqq 1),$$

where M is a constant independent of n and z.

Theorem 8.2 is a more favorable result than can be obtained by the standard method[27] from Theorem 8.1 by the use of Lagrange's interpolation formula on the segment $-1 \leqq z \leqq 1$ for interpolation by a polynomial $p_n(z)$ of degree n to a function $f(z)$ in the $n + 1$ roots of the Legendre polynomial $P_{n+1}(z)$:

$$p_n(z) = \sum_{k=1}^{n+1} f(z_k)l_k(z), \qquad \sum_{k=1}^{n+1} |l_k(z)| \leqq M_0 \cdot n \qquad (-1 \leqq z \leqq 1).$$

The conclusion of Theorem 8.2 holds if the $P_n(z)$ belong to a large class of Jacobi polynomials.

Again, Theorems 3.4, 7.1, and inequality (5.9) yield

THEOREM 8.3. *Let E and $f(z)$ satisfy the hypothesis of Theorem* 8.1, *and let $p_n(z)$ be the polynomial of degree n which interpolates to $f(z)$ in the $n + 1$ roots of the polynomial $Q_{n+1}(z)$, orthogonal on E with respect to a suitable weight function. Then we have*

$$\int_{-1}^{1} |f(z) - p_n(z)|^2 \, dz \leqq \frac{M}{n^{2p+2\alpha}\rho^{2n}}, \tag{8.1}$$

where M is independent of n and z.

We state now a final result on the segment, a consequence of Theorems 3.3 and 5.2, and inequality (7.5).

THEOREM 8.4. *Let E satisfy the hypothesis of Theorem* 8.1, *and let $f(z)$ be of class F_q on C_ρ. Let $p_n(z)$ be the polynomial of degree n which interpolates to $f(z)$ in the $n + 1$ roots of $T_{n+1}(z)$. Then we have*

$$|f(z) - p_n(z)| \leqq \frac{M}{n^{q-\frac{1}{2}}\rho^n} \qquad (-1 \leqq z \leqq 1),$$

where M is a constant independent of n and z.

9. **Lemniscates.**[28] Theorems 3.2, 3.3, and 3.4 apply at once to the case of lemniscates, by virtue of the inequalities of §6. In particular, Theorems 3.2, 7.1, and inequality (6.4) yield

THEOREM 9.1. *Let $p(z)$ be the polynomial $p(z) \equiv (z - \beta_1) \cdots (z - \beta_\lambda)$, and let Γ_μ denote generically the lemniscate $|p(z)| = \mu > 0$. Let Γ_{μ_1} consist of a finite number of mutually exterior analytic Jordan curves. Let $f(z)$ be of class $L(p, \alpha)$ $(0 < \alpha \leqq 1)$ on Γ_{μ_1}. Let $p_n(z)$ be the polynomial of degree n which interpolates to $f(z)$ in the $n + 1$ $(= s\lambda + r, 0 \leqq r < \lambda)$ points, namely, the $\beta_1, \beta_2, \cdots, \beta_r$*

[27] Shohat [1].

[28] Cf. Walsh and Curtiss [1], and also a forthcoming paper by Curtiss.

counted each $s + 1$ times and the points $\beta_{r+1}, \cdots, \beta_\lambda$ counted each s times. Then we have for each μ $(0 < \mu < \mu_1)$

$$(9.1) \qquad |f(z) - p_n(z)| \leqq \frac{M}{n^{p+\alpha}} \left(\frac{\mu}{\mu_1}\right)^{n/\lambda}, \qquad z \text{ on } \Gamma_\mu,$$

where M is independent of n and z.

We may here set $\rho = (\mu_1/\mu)^{1/\lambda}$, where ρ has the usual significance.

Of course the unit circle is a lemniscate with a single pole at the origin. It is of interest to note that in this case we have (9.1) on Γ_μ, although on $\Gamma_{\mu_1} = C_\rho$ we have only[29]

$$|f(z) - p_n(z)| \leqq \frac{M \log n}{n^{p+\alpha}}.$$

Chapter IV. Further results

10. Lipschitz conditions. Continuity properties of $f^{(p)}(z)$. We now proceed to discuss the relation between both Lipschitz conditions and derivatives on the *boundary* of a Jordan region, and the corresponding conditions in the *closed* regions. We shall make essential use of Theorem 10.1 below, which is in the main due to Warschawski [1]; in addition, the present writers have received orally some valuable suggestions from Dr. Warschawski concerning the material of the present section.

Let C be a Jordan curve of the z-plane, whose interior is denoted by R; let the set $R + C$ be denoted by $\bar{C}$. Let the function $f(z)$ be analytic in R, continuous in $\bar{C}$. If z_0 is a point of R, the derivative $f'(z)$ exists for the value $z = z_0$ and is continuous there. But if z_0 is a point of C, there are three different limits that can be considered appropriately: (i) the one-dimensional derivative of $f(z)$ on C is

$$(10.1) \qquad \lim_{z \to z_0,\, z \text{ on } C} \frac{f(z) - f(z_0)}{z - z_0},$$

provided this limit exists; (ii) the two-dimensional derivative of $f(z)$ is

$$(10.2) \qquad \lim_{z \to z_0,\, z \text{ in } \bar{C}} \frac{f(z) - f(z_0)}{z - z_0},$$

provided the limit exists; (iii) the derivative $f'(z)$ is continuous in the two-dimensional sense at $z = z_0$ if we have

$$(10.3) \qquad \lim_{z \to z_0,\, z \text{ in } \bar{C}} f'(z) = f'(z_0),$$

where $f'(z_0)$ is defined either in the sense (i) or (ii). The purpose of the present section is the study of relations of equality among these three limits (i), (ii), and (iii). We have thus far mentioned only $f'(z)$ but analogous questions obviously arise for higher derivatives.

[29] Compare Sewell [1].

THEOREM 10.1. *Let R be the interior of a Jordan curve C, let $R + C$ be denoted by $\bar{C}$, and let $f(z)$ be analytic in R, continuous in $\bar{C}$. Let α be fixed, with $0 < \alpha \leqq 1$, and suppose for all z and $z_0 \neq z$ on C we have*

$$\left| \frac{f(z) - f(z_0)}{(z - z_0)^\alpha} \right| \leqq L, \tag{10.4}$$

where L is independent of z_0 and z; then this inequality is valid for all z and z_0 in $\bar{C}$.

Theorem 10.1 is a slight generalization of a theorem due to Warschawski [1], namely, Theorem 10.1 with the hypothesis (10.4) for z fixed on C and z_0 arbitrary on C, and the conclusion (10.4) for this fixed z and for z_0 arbitrary in $\bar{C}$. In proving the generalization we shall need to employ Warschawski's form of the theorem.

Case I: $\alpha = 1$. For z_0 fixed in R, the function of z:

$$F_{z_0}(z) = \frac{f(z) - f(z_0)}{z - z_0}, z \neq z_0, \qquad F_{z_0}(z_0) = f'(z_0),$$

is bounded in the neighborhood of z_0, hence is analytic for z in R even for $z = z_0$, and is continuous for z in $\bar{C}$. But for z_0 in R and z on C we have (Warschawski)

$$| F_{z_0}(z) | \leqq L, \tag{10.5}$$

whence the inequality (10.5) holds for z_0 in R and z in $\bar{R}$. Thus inequality (10.4) is known for both z and z_0 on C, one of these points on C and the other in R, and for both points in R. The proof is complete in Case I.

Case II: $\alpha < 1$. For z_0 fixed in R, the function of z:

$$P_{z_0}(z) = \log \left| \frac{f(z) - f(z_0)}{(z - z_0)^\alpha} \right|$$

is single valued and harmonic throughout the interior of R except when $z = z_0$:

$$\lim_{z \to z_0} \frac{|f(z) - f(z_0)|}{|z - z_0|^\alpha} = 0,$$

and except at points z for which $f(z) = f(z_0)$; the latter set of points D has no limit point in R (we exclude the trivial case $f(z) \equiv f(z_0)$); at a point of D the function $P_{z_0}(z)$ becomes negatively infinite. If z in R approaches a point z_1 of C for which $f(z_1) = f(z_0)$, the function $P_{z_0}(z)$ becomes negatively infinite. If z in R approaches any other point z_1 of C, the function $P_{z_0}(z)$ is continuous there in the two-dimensional sense, and hence (Warschawski) its value is not greater than $\log L$. Then the superior limit of $P_{z_0}(z)$ as z in R approaches C is not greater than $\log L$, so by the maximum principle for harmonic functions we have $P_{z_0}(z) \leqq \log L$ for z in R whenever $P_{z_0}(z)$ is defined. Thus inequality (10.4) is known for both z and z_0 on C, for one of these points in R and the other on C, and for both of those points in R, with $0 < \alpha \leqq 1$. Theorem 10.1 is now completely established.

We shall employ Theorem 10.1 in the proof of

THEOREM 10.2. *Let R be the interior of a Jordan curve C, let $R + C$ be denoted by $\bar{C}$, and let $f(z)$ be analytic in R, continuous in $\bar{C}$. Suppose further that $f(z)$ has a derivative $f'(z)$ in the one-dimensional sense at each point z_0 of C. If*

$$\left| \frac{f(z) - f(z_0)}{z - z_0} \right| \tag{10.6}$$

is uniformly bounded for all z and z_0 on C, then for every z_0 on C we have

$$\lim_{z \to z_0,\, z \text{ in } \bar{C}} \frac{f(z) - f(z_0)}{z - z_0} = \lim_{z \to z_0,\, z \text{ on } C} \frac{f(z) - f(z_0)}{z - z_0} = f'(z_0);$$

that is to say, the two-dimensional derivative $f'(z_0)$ exists on C and is equal to the one-dimensional derivative $f'(z_0)$.

Theorem 10.2 follows at once from Theorem 10.1 and a well-known theorem due to Lindelöf, for if z_0 is fixed on C and z lies in $\bar{C}$, the function (10.6) is uniformly bounded (Warschawski); this function approaches one and the same limit when z on C approaches z_0 from either direction. Consequently the limit exists in the two-dimensional sense.

If we modify the hypothesis of Theorem 10.2 so that $f'(z)$ shall exist in the one-dimensional sense merely at a particular point z_0, and if the function (10.6) is uniformly bounded for this fixed z_0 and for z arbitrary on C, the proof and conclusion remain valid, by use of the theorems of Warschawski and Lindelöf.

Closely related to Theorems 10.1 and 10.2 is a result on boundedness of the difference-quotient:

THEOREM 10.3. *Let C be a rectifiable Jordan curve with the property that there exists a number $k > 1$ such that if r denotes the distance $|z_1 - z_2|$ between any two points of C and s denotes the shorter arc of C joining z_1 and z_2, then we have*

$$\frac{s}{r} \leqq k. \tag{10.7}$$

Let $f(z)$ be continuous on C and have a bounded derivative on C in the one-dimensional sense: $|f'(z)| \leqq M$. Then for z_1 and z_2 on C we have

$$\left| \frac{f(z_2) - f(z_1)}{z_2 - z_1} \right| \leqq Mk.$$

Theorem 10.3 follows at once by (4.27) (see Saks [1], p. 207).

Another theorem is complementary to Theorems 10.1 and 10.2:

THEOREM 10.4. *Let R be the interior of a Jordan curve C, let $R + C$ be denoted by $\bar{C}$, and let $f(z)$ be analytic in R, continuous in $\bar{C}$. Let $f(z)$ have a continuous derivative $f'(z)$ in the one-dimensional sense on C, and suppose there exists a constant M such that*

$$(10.8) \qquad \left| \frac{f(z_2) - f(z_1)}{z_2 - z_1} \right| \leqq M,$$

for arbitrary z_1 and z_2 on C. Then (by Theorem 10.2) *the one-dimensional derivative $f'(z)$ on C is also a two-dimensional derivative. Moreover, we have for z_0 on C*

$$\lim_{z \to z_0,\, z \text{ in } \bar{C}} f'(z) = f'(z_0);$$

that is to say, the function $f'(z)$ is continuous in the two-dimensional sense on C, hence continuous in the two-dimensional sense in $\bar{C}$.

Theorem 10.3 shows that the inequality (10.8) is here a consequence of inequality (10.7) for the curve C.

Let the function $z = \lambda(w)$ map $|w| < 1$ conformally onto R. Consider for a fixed positive δ the function

$$(10.9) \qquad F(w, \delta) \equiv \frac{f[\lambda(e^{i\delta} w)] - f[\lambda(w)]}{\lambda(e^{i\delta} w) - \lambda(w)},$$

which is analytic in $|w| < 1$, even in the point $w = 0$ when suitably defined there, and is continuous in $|w| \leqq 1$. Then $F(w, \delta)$ can in $|w| < 1$ be represented by the Poisson integral:

$$(10.10) \qquad F(w, \delta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} F(e^{i\theta}, \delta) \frac{(1 - r^2)\, d\theta}{1 + r^2 - 2r \cos(\theta - \alpha)}, \qquad w = re^{i\alpha}.$$

From (10.9) it follows that for fixed w in $|w| < 1$, we have

$$(10.11) \qquad \lim_{\delta \to 0} F(w, \delta) = f'(z)]_{z=\lambda(w)};$$

this equation is also valid for fixed w on $|w| = 1$, where the second member now indicates one-dimensional derivative taken along C:

$$(10.12) \qquad \lim_{\delta \to 0} F(e^{i\theta}, \delta) = f'[\lambda(e^{i\theta})].$$

This latter function in (10.12) is known to be continuous on $|w| = 1$.

In equation (10.10) we keep w fixed, $|w| < 1$, and allow δ to approach zero. The function $F(e^{i\theta}, \delta)$ is by (10.8) bounded uniformly for all δ and θ. It follows by the Osgood-Lebesgue theorem that we may take the limit in (10.10) under the integral sign as δ approaches zero, making use of (10.11) and (10.12):

$$(10.13) \qquad f'[\lambda(w)] = \frac{1}{2\pi} \int_{-\pi}^{\pi} f'[\lambda(e^{i\theta})] \frac{(1 - r^2)\, d\theta}{1 + r^2 - 2r \cos(\theta - \alpha)}, \qquad w = re^{i\alpha}.$$

Equation (10.13) is the representation of a function by Poisson's integral, the function $f'[\lambda(e^{i\theta})]$ being a continuous function of θ. It is then well known that these functional values, continuous with respect to θ, are taken on continuously in the two-dimensional sense by the first member of (10.13):

$$\lim_{w \to e^{i\theta}, |w|<1} f'[\lambda(w)] = f'[\lambda(e^{i\theta})],$$

or for z_0 on C,

$$\lim_{z \to z_0,\, z \text{ in } R} f'(z) = f'(z_0);$$

the corresponding equation holds for z on C, so Theorem 10.4 is established.

The special case of Theorem 10.4 in which C is assumed analytic (here inequality (10.8) is always valid, by Theorem 10.3) has recently been proved by the present writers [2]. The special cases of Theorem 10.4 where C is the unit circle had been previously established by Hardy and Littlewood [1].

The following result is an immediate consequence of Theorems 10.4 and 10.1:

THEOREM 10.5. *Let R be the interior of a Jordan curve C, let $R + C$ be denoted by $\bar{C}$, and let $f(z)$ be analytic in R, continuous in $\bar{C}$. Let the p-th derivative of $f(z)$ exist in the one-dimensional sense on C and be continuous on C. Also let*

$$\left| \frac{f^{(k)}(z) - f^{(k)}(z_0)}{z - z_0} \right|, \qquad z \text{ and } z_0 \text{ on } C,\ k = 0, 1, 2, \cdots, p-1, \tag{10.14}$$

be uniformly bounded. Then the two-dimensional derivatives $f'(z), f''(z), \cdots, f^{(p)}(z)$ exist on C and are continuous in the two-dimensional sense in $\bar{C}$. If $f^{(p)}(z)$ satisfies a Lipschitz condition of order α on C $(0 < \alpha \leqq 1)$, then $f^{(p)}(z)$ satisfies this same Lipschitz condition in $\bar{C}$.

Under mild conditions on C, the uniform boundedness of (10.14) for $k = 0, 1, 2, \cdots, p-1$ is a consequence of the existence and continuity on C of $f^{(p)}(z)$, by Theorem 10.3.

Chapter V. Further problems

11. **Problem γ.** In §1 we defined Problem γ as the study of the convergence on C_ρ or $C_{\rho'}$ of a sequence of polynomials whose degree of convergence is given on the closed limited set E whose boundary is C.

Suppose for definiteness that the complement of E is connected and regular, and let us assume

$$|f(z) - p_n(z)| \leqq \epsilon_n, \qquad z \text{ on } E,\ n = 1, 2, \cdots, \tag{11.1}$$

where $\epsilon_{n+1} \leqq \epsilon_n$ and where $p_n(z)$ is a polynomial of degree n. Then we have also

$$|p_{n+1}(z) - p_n(z)| \leqq 2\epsilon_n, \qquad z \text{ on } E,$$

whence by the generalized Bernstein lemma (Walsh [1], pp. 77–78),

$$|p_{n+1}(z) - p_n(z)| \leqq 2\epsilon_n \nu^{n+1}, \qquad z \text{ on } C_\nu,\ \nu > 1. \tag{11.2}$$

If ϵ_n approaches zero as n becomes infinite, we have uniformly for z on E

$$f(z) = p_1(z) + [p_2(z) - p_1(z)] + [p_3(z) - p_2(z)] + \cdots. \tag{11.3}$$

If $\sum \epsilon_n \nu^{n+1}$ converges, we may consider the second member of (11.3) to define $f(z)$ on C_ν, hence also throughout the closed interior of C_ν. By (11.2) we have

$$|f(z) - p_n(z)| \leqq 2 \sum_{k=n}^{\infty} \epsilon_k \nu^{k+1}, \qquad z \text{ on } C_\nu. \tag{11.4}$$

The method just indicated can be applied to all the sequences of polynomials studied in §§7-9 when the degree of convergence on E is known. It is not essential that the complement of E be simply connected. We formulate explicitly a consequence of Theorem 7.2:

THEOREM 11.1. *Let C and C_ρ be contours, let $f(z)$ be of class $L(p, \alpha)$ $(0 < \alpha \leqq 1)$ on C_ρ. Let $p_n(z)$ be the unique polynomial of degree n which interpolates to $f(z)$ in $n + 1$ equally distributed points of C. Then if $p \geqq 1$, we have*

$$|f(z) - p_n(z)| \leqq \frac{M_1}{n^{p+\alpha-1}}, \qquad z \text{ on } C_\rho, \tag{11.5}$$

where M_1 is independent of n and z.

There is a similar consequence of Theorem 7.3, which in the notation of that theorem is

THEOREM 11.2. *Let C be a Jordan curve, let $f(z)$ be of class $L(p, \alpha)$ $(0 < \alpha \leqq 1)$ on C_ρ. Let $p_n(z)$ be the polynomial of degree n which interpolates to $f(z)$ in the $n + 1$ roots of $T_{n+1}(z)$, where C is analytic, or in the $n + 1$ roots of $F_{n+1}(z)$, where C is of type A, or in the $n + 1$ roots of $Q_{n+1}(z)$, where C is analytic. Then if $p \geqq 1$ we have* (11.5) *satisfied, where M_1 is independent of n and z.*

Inequality (11.4) was derived under the hypothesis (11.1), concerning the modulus on E of the *maximum* deviation of $p_n(z)$ from $f(z)$. If approximation is measured by an integral of a power of the deviation, a similar inequality can be obtained (compare Sewell [4]). Let C be a rectifiable Jordan curve, let $p_n(z)$ be a polynomial of degree n, and suppose $(m > 0)$

$$\int_C |f(z) - p_n(z)|^m |dz| \leqq \epsilon_n^m \qquad (n = 0, 1, 2, \cdots), \tag{11.6}$$

where $\epsilon_{n+1} \leqq \epsilon_n$. By general algebraic inequalities (Walsh [1], p. 93) we have

$$\int_C |p_{n+1}(z) - p_n(z)|^m |dz| \leqq N_1 \epsilon_n^m,$$

where N_1 is independent of n. By an integral analogue of Bernstein's lemma (Walsh [1], p. 92) we now have

$$|p_{n+1}(z) - p_n(z)| \leqq N\epsilon_n \nu^{n+1}, \qquad z \text{ on } C_\nu, \nu > 1,$$

where N is independent of n and z. If $\sum \epsilon_n \nu^{n+1}$ converges, we may as before use (11.3) to define $f(z)$ on C_ν, hence also on and within C_ν. The function thus defined must coincide with $f(z)$ on C almost everywhere, for the integral of the m-th power of the absolute value of the difference is zero. Thus we have

$$(11.7) \qquad |f(z) - p_n(z)| \leqq N \sum_{k=n}^{\infty} \epsilon_k \nu^{k+1}, \qquad z \text{ on } C_\nu.$$

As an application of (11.7) as a consequence of (11.6) we state from Theorem 7.4:

THEOREM 11.3. *Let C be an analytic Jordan curve, let $f(z)$ be of class $L(p, \alpha)$ $(0 < \alpha \leqq 1)$ on C_ρ. Let $p_n(z)$ denote the polynomial of degree n which interpolates to $f(z)$ in the $n + 1$ roots of the polynomial $Q_{n+1}(z)$ of the set normal and orthogonal on C with respect to a positive and continuous weight function. Then if $p \geqq 1$ we have (11.5) satisfied, where M_1 is independent of n and z.*

It is to be noted that the number ν used in deriving (11.4) as a consequence of (11.1) and in deriving (11.7) as a consequence of (11.6) requires no knowledge of the number ρ defined in terms of the function $f(z)$, namely, ρ is the largest ρ' such that $f(z)$ is analytic throughout the interior of $C_{\rho'}$; the number ν is entirely arbitrary, subject to the inequality $\nu > 1$, and can be employed with a given sequence $p_n(z)$ whenever $\sum \epsilon_n \nu^{n+1}$ converges. For instance, the method of proof already given shows that *under the conditions of Theorems* 11.1, 11.2, *or* 11.3 *we have*

$$(11.8) \qquad |f(z) - p_n(z)| \leqq \frac{M' \nu^n}{n^{p+\alpha} \rho^n}, \qquad z \text{ on } C_\nu,\ 1 < \nu < \rho,$$

where M' is independent of n and z.

This last remark can also be established by the results of §3. For instance under the hypothesis of Theorem 4.7 we have by the conclusion of that theorem and by Bernstein's lemma the inequality

$$(11.9) \qquad \left| \frac{\omega_n(z)}{\omega_n(t)} \right| \leqq \frac{M_0' \nu^n}{\rho^n}, \qquad z \text{ on } C_\nu,\ t \text{ on } C_\rho,$$

where M_0' is independent of n, t, and z. Under the hypothesis of Theorem 11.1, Theorems 7.1 and 3.2 (in which E is now taken as the closed interior of C_ν) then yield (11.8). In proving (11.8) under the conditions of Theorem 11.3, we make further use of the lemma on integrals of polynomials (Walsh [1], p. 92), applied now to $Q_{n+1}(z)$.

Still other results on Problem γ are to be established in §12.

12. **Best approximation. Examples.** If E is a closed limited point set on which the function $f(z)$ is continuous, the *polynomial $K_n(z)$ of degree n of best approximation to $f(z)$ on E in the sense of Tchebycheff* is that polynomial (known to exist and to be unique if E contains at least $n + 1$ points) of degree n for which

$$\max [|f(z) - K_n(z)|, z \text{ on } E]$$

is least. If we have at hand a specific polynomial $p_n(z)$ of degree n such that

$$|f(z) - p_n(z)| \leqq \epsilon_n, \qquad z \text{ on } E,$$

then we obviously have

$$|f(z) - K_n(z)| \leqq \epsilon_n, \qquad z \text{ on } E. \tag{12.1}$$

This remark enables us to state some immediate results on the convergence of the sequence $K_n(z)$.

THEOREM 12.1. *Let C be a contour, a finite line segment, or a lemniscate, and let C_ρ consist of a finite number of mutually exterior Jordan curves. Let $f(z)$ be of class $L(p, \alpha)$ $(0 < \alpha \leqq 1)$ on C_ρ.*

If $K_n(z)$ denotes the polynomial of degree n of best approximation to $f(z)$ on C_ρ in the sense of Tchebycheff, we have

$$|f(z) - K_n(z)| \leqq \frac{M}{n^{p+\alpha}}, \qquad z \text{ on } C_\rho. \tag{12.2}$$

If $L_n(z)$ denotes the polynomial of degree n of best approximation to $f(z)$ on C in the sense of Tchebycheff, we have

$$|f(z) - L_n(z)| \leqq \frac{M_1}{n^{p+\alpha}\rho^n}, \qquad z \text{ on } C, \tag{12.3}$$

$$|f(z) - L_n(z)| \leqq \frac{M_2 \nu^n}{n^{p+\alpha}\rho^n}, \qquad z \text{ on } C_\nu,\ 1 < \nu < \rho, \tag{12.4}$$

$$|f(z) - L_n(z)| \leqq \frac{M_3}{n^{p+\alpha-1}}, \qquad z \text{ on } C_\rho, \tag{12.5}$$

where the validity of (12.5) requires $p > 1$. The numbers M, M_1, M_2, M_3 are all constants independent of n and z.

Inequality (12.2) follows from Theorem 7.1, due to Curtiss. Inequality (12.3) follows from Theorems 7.2, 8.1, and 9.1, under the respective conditions on C; inequality (12.3) is valid if C is a lemniscate, even if that lemniscate has multiple points. Inequality (12.4) is a special case of (11.8), and follows from (12.3) just as (11.4) follows from (11.1). Inequality (12.5) is a special case of (11.4) and is similarly proved from (12.3).

Instead of studying approximation in the sense of Tchebycheff, we may also study approximation in the sense of least m-th powers, with $m > 0$. Polynomials of best approximation of prescribed degree always exist, and are unique if $m > 1$. The method of proof by comparison is useful in obtaining the relation analogous to (12.1); we derive the inequality

$$\int_C |f(z) - K_n(z)|^m \, |dz| \leqq \epsilon_n, \tag{12.6}$$

where $K_n(z)$ now denotes the (or a) polynomial of degree n of best approximation in the sense of least m-th powers, and where we assume the existence of a comparison sequence of polynomials $p_n(z)$ of respective degrees n with the property

$$\int_C |f(z) - p_n(z)|^m |dz| \leqq \epsilon_n.$$

We shall establish the

COROLLARY. *Let C, and C_ρ, and $f(z)$ satisfy the conditions of Theorem* 12.1. *Let m be positive.*

If $K_n(z)$ denotes the polynomial of degree n of best approximation to $f(z)$ on C_ρ in the sense of least m-th powers, then we have

$$\int_{C_\rho} |f(z) - K_n(z)|^m |dz| \leqq \frac{M}{n^{m(p+\alpha)}}. \tag{12.7}$$

If $L_n(z)$ denotes the polynomial of degree n of best approximation to $f(z)$ on C in the sense of least m-th powers, then we have

$$\int_C |f(z) - L_n(z)|^m |dz| \leqq \frac{M_1}{n^{m(p+\alpha)}\rho^{mn}}, \tag{12.8}$$

$$|f(z) - L_n(z)| \leqq \frac{M_2 \nu^n}{n^{p+\alpha}\rho^n}, \qquad z \text{ on } C_\nu,\ 1 < \nu < \rho, \tag{12.9}$$

$$|f(z) - L_n(z)| \leqq \frac{M_3}{n^{p+\alpha-1}}, \qquad z \text{ on } C_\rho, \tag{12.10}$$

where the validity of (12.10) *requires $p > 1$. The numbers M, M_1, M_2, M_3 are all constants independent of n and z.*

Inequality (12.7) is a consequence of (12.2); inequality (12.8) is a consequence of (12.3); inequalities (12.9) and (12.10) are special cases of (11.7), in the light of (12.8) in the rôle of (11.6). However, the integral analogue of Bernstein's Lemma as hitherto formulated (Walsh [1], p. 92) applies merely to a *single* Jordan arc or curve C; the proof of (11.7) for the present purposes requires a new lemma which contemplates the case that C consists of several components; the original proof of that analogue does not generalize directly to the present situation:

LEMMA 12.1. *Let C consist of a finite number of mutually exterior analytic Jordan curves, and let $p > 0$ be arbitrary. There exists a number L' such that the inequality*

$$\int_C |P_n(z)|^p |dz| \leqq L^p,$$

where $P_n(z)$ is a polynomial in z of degree n, implies

$$|P_n(z)| \leqq L'LR^n, \qquad z \text{ on } C_R,$$

where L' depends on C and R but not on n or $P_n(z)$.

As usual we denote by K the complement with respect to the extended plane of the closed interior of C, and by $\varphi(z)$ the mapping function for K. The func-

tion $| P_n(z)/[\varphi(z)]^n |^p$ is single valued and continuous throughout the closure of the region K of the extended plane, even at infinity if suitably defined there, and is subharmonic in K. This function is not greater than the function which coincides with it on C and is continuous on $K + C$, harmonic in K. Consequently we have for z in K

$$\left| \frac{P_n(z)}{[\varphi(z)]^n} \right|^p \leqq \frac{1}{2\pi} \int_C \left| \frac{P_n(z')}{[\varphi(z')]^n} \right|^p \frac{\partial G}{\partial n}\, ds = \frac{1}{2\pi} \int_C | P_n(z') |^p \frac{\partial G}{\partial n}\, ds,$$

where $G(z, z')$ denotes Green's function for K with pole in the point z and z' is the running coördinate on C. For z on C_R, the function $\partial G/\partial n$ on C is uniformly bounded, say not greater than L_1 in absolute value, so we may write from our hypothesis

$$\left| \frac{P_n(z)}{[\varphi(z)]^n} \right|^p \leqq \frac{1}{2\pi} L_1 L^p, \qquad z \text{ on } C_R.$$

For z on C_R we have $| \varphi(z) | = R$, so the conclusion of Lemma 12.1 follows at once.

The restriction that C shall be composed of *analytic* Jordan curves is a matter only of simplicity. Let us suppose for instance that C is composed of a finite number of rectifiable Jordan arcs or curves with the property that

$$\int_C \left| \frac{\partial G}{\partial n} \right|^{1/(1-\alpha)} ds, \qquad\qquad 0 < \alpha < 1,$$

exists and is uniformly bounded, say not greater than L_2, for z on C_R. Then we may write by the Hölder inequality (3.8) and by the subharmonic property of the function $| P_n(z)/[\varphi(z)]^n |^{\alpha p}$,

$$2\pi \left| \frac{P_n(z)}{[\varphi(z)]^n} \right|^{\alpha p} \leqq \int_C \left| \frac{P_n(z')}{[\varphi(z')]^n} \right|^{\alpha p} \frac{\partial G}{\partial n}\, ds \leqq \left(\int_C \left| \frac{P_n(z')}{[\varphi(z')]^n} \right|^p ds \right)^{\alpha} \left(\int_C \left| \frac{\partial G}{\partial n} \right|^{1/(1-\alpha)} ds \right)^{1-\alpha}$$
$$\leqq L^{\alpha p} \cdot L_2^{1-\alpha}.$$

The previous reasoning is valid in establishing the conclusion of the lemma.

The behavior of $\partial G/\partial n$ on C, with respect to continuity, integrability, etc., is entirely similar to the behavior on C of $\partial G_1/\partial n$, where G_1 now indicates Green's function for the exterior of the various single components of C, with pole at infinity. For if the components of C are denoted by $C_1, C_2, \cdots, C_\nu$, we map successively onto the exterior of a circle (so that the points at infinity correspond to each other) the exterior of C_1, the exterior of the transform of C_2, the exterior of the last transform of C_3, and so on. Each of these transformations but one is analytic on each C_k. The last map provides a transformation of K onto a region bounded by a finite number of analytic Jordan curves; in this latter situation $\partial G/\partial n$ is analytic on the transform of C.

Of course the restriction on $\partial G/\partial n$ or even the requirement that C be rectifiable may be omitted in Lemma 12.1 provided the original inequality is replaced by the inequality

$$\int_C |P_n(z)|^p \, du \leqq L^p,$$

in the notation of §4.

If C is a contour, the function $\partial G/\partial n$ is continuous on C and is uniformly bounded for z on C_R ; if C is a lemniscate or line segment, some power greater than unity of $\partial G/\partial n$ is integrable on C and the integral is uniformly bounded for z on C_R ; in either case the remarks supplementary to Lemma 12.1 apply; the conclusion of the corollary follows.

Analogues of Theorem 12.1 and its corollary, when we now require that $f(z)$ shall be of class F_q on C_ρ instead of requiring a Lipschitz condition on $f^{(p)}(z)$, can be proved at once and are left to the reader.

For the special case that C is a circle or line segment, Theorem 12.1 and its corollary can be established by relatively elementary methods, as can also Theorem 12.2 below and its corollary (see Walsh and Sewell [3, 4]).

For comparison and contrast with Theorem 12.1 we state a result (Walsh and Sewell [2]) which concerns both Problem β and Problem γ, and is proved by the method used in proving (11.4):

THEOREM 12.2. *Let E with boundary C be a closed limited point set whose complement K is connected, and is regular in the sense that there exists a function $w = \varphi(z)$ which maps K conformally but not necessarily uniformly onto $|w| > 1$ so that the points at infinity correspond to each other. Let the locus C_ρ consist of a finite number of mutually exterior analytic Jordan curves. Let $f(z)$ be defined on E, and for $n = 1, 2, \cdots$ let a polynomial $p_n(z)$ of degree n exist such that*

$$|f(z) - p_n(z)| \leqq \frac{M}{n^{p+\alpha+1}\rho^n}, \qquad z \text{ on } E,\ 0 < \alpha \leqq 1, \tag{12.11}$$

where M is a constant independent of n and z, and where p is a non-negative integer. Then when suitably defined exterior to E, the function $f(z)$ is analytic in C_ρ and continuous in the corresponding closed region. We have the inequalities

$$|f(z) - p_n(z)| \leqq \frac{M_1 \nu^n}{n^{p+\alpha+1}\rho^n}, \qquad z \text{ on } C_\nu,\ 1 < \nu < \rho, \tag{12.12}$$

$$|f(z) - p_n(z)| \leqq \frac{M_2}{n^{p+\alpha}}, \qquad z \text{ on } C_\rho; \tag{12.13}$$

a consequence of (12.13) is that the p-th derivative $f^{(p)}(z)$ exists on C_ρ in the one-dimensional sense and satisfies there the condition, for $|z_1 - z_2|$ sufficiently small,

$$|f^{(p)}(z_1) - f^{(p)}(z_2)| \leqq L\,|z_1 - z_2|^\alpha \cdot |\log|z_1 - z_2||^\beta, \quad z_1 \text{ and } z_2 \text{ on } C_\rho, \tag{12.14}$$

where $\beta = 0$ if $\alpha < 1$ and $\beta = 1$ if $\alpha = 1$, and L is a constant independent of z_1 and z_2.

Comparison of Theorems 12.1 and 12.2 shows that they are not exact converses

each of the other, for there is a discrepancy of unity in the exponents of (12.3) and (12.11). *This discrepancy is inherent in the nature of the problem,* as we shall now proceed to show. Expressed in other words, we shall show that *if we restrict ourselves to degree of convergence on C or E as measured by such an expression as $M/n^{\delta}\rho^{n}$, then for arbitrary α $(0 < \alpha \leqq 1)$ the inequalities* (12.3) *and* (12.11) *in Theorems* 12.1 *and* 12.2 *cannot be improved.*

Let us choose

$$
\begin{aligned}
f(z) &\equiv -\int_0^z \log(1-z)\,dz = \int_0^z \left(z + \frac{z^2}{2} + \frac{z^3}{3} + \cdots\right) dz \\
&= \frac{z^2}{2} + \frac{z^3}{2\cdot 3} + \frac{z^4}{3\cdot 4} + \cdots, \qquad |z| < 1, \\
&\equiv (1-z)\log(1-z) + z.
\end{aligned} \tag{12.15}
$$

If C is the circle $|z| = \frac{1}{2}$, and if $p_n(z)$ represents the partial sum of degree n of the Taylor development of $f(z)$, we have

$$
|f(z) - p_n(z)| \leqq \sum_{k=n}^{\infty} \frac{1}{k(k+1)2^{k+1}} \leqq \frac{1}{n^2} \sum_{k=n}^{\infty} \frac{1}{2^{k+1}} = \frac{1}{n^2 2^n}. \tag{12.16}
$$

Inequality (12.16) is precisely of the form (12.11), with $\rho = 2$, $p = 0$, $\alpha = 1$. It will be noted that $f'(z)$ is not continuous on $C_\rho : |z| = 1$, and that $f(z)$ does not satisfy a Lipschitz condition on C_ρ of order unity, but rather a condition of form (12.14) with $\alpha = 1$. Thus Theorem 12.2 for the case $p = 0$, $\alpha = 1$, cannot be improved, in the sense mentioned.

For the case $p = 0$, $0 < \alpha < 1$, we replace (12.15) by the equation

$$
f(z) \equiv (1-z)^\alpha \equiv \sum_{k=0}^{\infty} a_k z^k, \qquad |z| < 1,
$$

from which it follows (Dienes [1], p. 24) that we have

$$
|a_n| \leqq \frac{M}{n^{\alpha+1}}.
$$

Consequently we may write for z on C: $|z| = \frac{1}{2}$

$$
\left|f(z) - \sum_{k=0}^{n} a_k z^k\right| \leqq \sum_{k=n+1}^{\infty} \frac{M}{2^k k^{\alpha+1}} \leqq \frac{1}{n^{\alpha+1}} \sum_{k=n+1}^{\infty} \frac{M}{2^k} = \frac{M}{2^n n^{\alpha+1}}.
$$

This inequality is of the form (12.11), with $\rho = 2$, $p = 0$; it will be noted that $f(z)$ is continuous for $|z| \leqq 1$ and satisfies on $C_\rho : |z| = 1$ a Lipschitz condition of order α but of no higher order. Thus Theorem 12.2 for the case $p = 0$, $0 < \alpha < 1$, cannot be improved, in the sense mentioned.

For arbitrary positive integral p and $0 < \alpha \leqq 1$ it is also true that Theorem 12.2 cannot be improved in this same sense; suitable counter-examples (Gegen-

beispiele) are obtained by taking successive integrals of the functions just studied.[30]

As a corresponding example in the opposite direction, we take the illustration (similar to one used by Hardy and Littlewood [1] for another purpose)

$$f(z) \equiv \sum_{k=0}^{\infty} \frac{z^{n_k}}{2^k}, \qquad n_{k+1} > n_k, \tag{12.17}$$

which is analytic for $|z| < 1$, continuous for $|z| \leqq 1$; the integers n_k are to be further restricted. Let α be arbitrary, with $0 < \alpha < 1$. We define n_{k+1} for k sufficiently large by the inequality[31]

$$2^{(k-1)/\alpha} < n_{k+1} - 1 \leqq 2^{k/\alpha};$$

this definition is consistent with the requirement $n_{k+1} > n_k$ for all k, for the inequality $k + 1 < 2^{(k-1)/\alpha}$ for k sufficiently large follows from the inequality

$$\frac{\log (k+1)}{k-1} < \frac{1}{\alpha} \log 2,$$

which is satisfied for k sufficiently large. Moreover, the relation

$$2^{k/\alpha} - 2^{(k-1)/\alpha} > 1$$

is satisfied for k sufficiently large, and is merely the equivalent of the inequality

$$2^{k-1} > \frac{1}{(2^{1/\alpha} - 1)^{\alpha}}.$$

Whenever we have $n_k \leqq n < n_{k+1}$, the partial sum $s_n(z)$ of degree n of the Taylor development of $f(z)$ is

$$\sum_{j=0}^{k} \frac{z^{n_j}}{2^j},$$

[30] Whenever the function $f(z) \equiv \sum a_k z^k$ is analytic for $|z| \leqq r$, and the numbers a_k are non-negative, we obviously have

$$\max [|f'(z) - \sum_{k=1}^{n} k a_k z^{k-1}|, |z| = r] = \sum_{k=n+1}^{\infty} k a_k r^{k-1}$$

$$\geqq n \sum_{k=n+1}^{\infty} a_k r^{k-1} = n \cdot \max [|f(z) - \sum_{k=1}^{n} a_k z^k|, |z| = r]/r.$$

Whenever the function $f(z) \equiv \sum a_k z^k$ is analytic for $|z| \leqq r$ we also have

$$\int_{|z|=r} |f'(z) - \sum_{k=1}^{n} k a_k z^{k-1}|^2 \cdot |dz| = 2\pi r \sum_{k=n+1}^{\infty} k^2 |a_k|^2 r^{2k-2}$$

$$\geqq 2\pi n^2 r \sum_{k=n+1}^{\infty} |a_k|^2 r^{2k-2} = \frac{n^2}{r^2} \int_{|z|=r} |f(z) - \sum_{k=1}^{n} a_k z^k|^2 \cdot |dz|;$$

these two integrals represent the measure of *best* approximation on $|z| = r$ in the sense of least squares to the respective functions $f'(z)$ and $f(z)$.

whence we have on $|z| = 1$ for k sufficiently large

$$|f(z) - s_n(z)| \leqq \frac{1}{2^k} \leqq \frac{1}{(n_{k+1} - 1)^\alpha} \leqq \frac{1}{n^\alpha}.$$

It follows (Walsh and Sewell [2]) that $f(z)$ satisfies on $|z| = 1$ a Lipschitz condition of order α. Choose now for definiteness $\rho = 2$, let C be $|z| = 1/\rho$, and let $L_n(z)$ be the polynomial of degree n of best approximation to $f(z)$ on C in the sense of Tchebycheff. We set

$$\epsilon_n = \max [|f(z) - L_n(z)|, z \text{ on } C].$$

The polynomial $s_n(z)$ is the polynomial of degree n of best approximation to $f(z)$ on C in the sense of least squares, so we have

$$\pi\epsilon_n^2 \geqq \int_C |f(z) - L_n(z)|^2 |dz| \geqq \int_C |f(z) - s_n(z)|^2 |dz|.$$

For the particular value $n = n_{k+1} - 1$ this yields by (12.17) for k sufficiently large

$$\epsilon_n^2 \geqq \sum_{j=k+1}^{\infty} \frac{1}{2^{2n_j}\cdot 2^{2j}} > \frac{1}{2^{2n_{k+1}}\cdot 2^{2k+2}} = \frac{1}{2^6\cdot 2^{2n_{k+1}-2}\cdot 2^{2k-2}},$$

(12.18) $$\epsilon_n > \frac{1}{8\cdot 2^n\cdot n^\alpha}.$$

That is to say, inequality (12.18) obtains for an infinite sequence of indices n, so for the function $f(z)$ defined by (12.17) it is not possible to replace (12.3) by any inequality of the form

$$|f(z) - L_n(z)| \leqq \frac{M_1}{n^{\alpha'}\rho^n}, \qquad z \text{ on } C,$$

with $\alpha' > \alpha$, or even by any inequality of the form

$$|f(z) - L_n(z)| \leqq \frac{\eta(n)}{\rho^n\cdot n^\alpha}, \qquad z \text{ on } C,$$

where $\eta(n)$ approaches zero as n becomes infinite. Thus Theorem 12.1 for the case $p = 0$, $0 < \alpha < 1$, cannot be improved, in the sense mentioned.

For class $L(0, 1)$ Theorem 12.1 cannot be improved in this same sense, for suppose it could be shown that for every function of class $L(0, 1)$ on C_ρ the second member of (12.3) could be replaced by $M_1\rho^{-n}n^{-\beta}$, where β is some number greater than unity; we shall reach a contradiction. For the integral $F(z)$ of the special function $f(z)$ considered above, with $0 < \alpha < \beta - 1$,

$$F(z) \equiv \sum_{j=0}^{\infty} \frac{z^{n_j+1}}{2^j(n_j + 1)},$$

[21] We remark incidentally that it follows from Hadamard's gap theorem that here $f(z)$ cannot be continued analytically beyond the circle $|z| = 1$.

which belongs to class $L(0, 1)$, we have for the Tchebycheff polynomial $L_n(z)$ for $F(z)$ on C: $|z| = \frac{1}{2}$ by the kind of reasoning already given

$$\max [|F(z) - L_n(z)|, z \text{ on } C] \geqq \frac{m}{\rho^n n^{1+\alpha}},$$

where m is positive and independent of n; thus for this particular function $F(z)$ of class $L(0, 1)$ the second member of (12.3) cannot be replaced by $M_1\rho^{-n}n^{-\beta}$. Consequently Theorem 12.1 cannot be improved, in the sense considered, for the case $p = 0$, $\alpha = 1$. This remark then holds for $0 < \alpha \leqq 1$; it holds for arbitrary positive p and $0 < \alpha \leqq 1$ by replacing the specific functions $f(z)$ already used by successive indefinite integrals of those functions.

The corollary to Theorem 12.1 modifies that theorem by introducing integral measures of approximation instead of the Tchebycheff measure of approximation. There is a similar corollary to Theorem 12.2:

COROLLARY. *Let E be a closed limited point set whose boundary C is a contour, and let the complement K of E be connected. Let the function $f(z)$ be defined on E, and for $n = 1, 2, \cdots$ let a polynomial $p_n(z)$ exist such that we have* ($m > 0$)

$$\int_C |f(z) - p_n(z)|^m \, |dz| \leqq \frac{M}{n^{m(p+\alpha+1)}\rho^{mn}}, \qquad z \text{ on } E,\ 0 < \alpha \leqq 1, \tag{12.19}$$

where M is a constant independent of n and z, and where p is a non-negative integer. Then when suitably defined or redefined, the function $f(z)$ (originally determined by (12.19) *on C only almost everywhere) is analytic interior to C_ρ and continuous in the corresponding closed region. We have the inequalities* (12.12) *and* (12.13); *a consequence of the latter inequality is that $f^{(p)}(z)$ exists on C_ρ in the one-dimensional sense and satisfies there the inequality* (12.14).

This corollary is established by interpreting (12.12) and (12.13) as special cases of (11.7), proved from the remarks on Lemma 12.1; the consequence of (12.13) follows from a result previously published[32] (Walsh and Sewell [2]).

It is of course obvious that the conclusion (11.7) as well as the conclusions of the corollaries to Theorems 12.1 and 12.2 are not affected if we measure approximation by the integral of m-th powers *with a positive continuous norm function.*

In the sense in which inequalities (12.3) and (12.11) cannot be improved in Theorems 12.1 and 12.2, it is also true that the corresponding inequalities (12.8) and (12.19) of their corollaries cannot be improved, at least for the case $m = 2$. This fact can be shown by use of the examples already employed, by essentially the methods already used. The fact that these various inequalities cannot be improved has been previously stated in the literature (Walsh and Sewell [3, 4]); the statement is hereby justified.

We return to the general topic of Theorem 12.1. Let C_ρ be for definiteness the unit circle. Let the function $f(z)$ be of the form

[32] We need not assume here that C_ρ is a contour, by virtue of the remark made in connection with Theorem 7.2.

$$f(z) \equiv (z - z_0)^{p+\alpha} + f_1(z), \tag{12.20}$$

where z_0 lies on C_ρ and $f_1(z)$ is analytic in the closed interior of C_ρ, and where p is integral, $0 < \alpha \leqq 1$. It may be proved (Dienes [1], p. 24) that the coefficients a_n in the Taylor development $\sum a_n z^n$ of $f(z)$ satisfy the inequality $|a_n| \leqq M/n^{p+\alpha+1}$, and hence that we have for the polynomials $L_n(z)$ of degree n of best approximation to $f(z)$ on C: $|z| = 1/\rho < 1$ in the sense of Tchebycheff

$$\max [|f(z) - L_n(z)|, z \text{ on } C] \leqq \max [|f(z) - \sum_{k=0}^{n} a_k z^k|, z \text{ on } C]$$

$$= \max [|\sum_{k=n+1}^{\infty} \alpha_k z^k|, z \text{ on } C] \tag{12.21}$$

$$\leqq \frac{M_1}{n^{p+\alpha+1}} \sum_{k=n+1}^{\infty} \frac{1}{\rho^k} \leqq \frac{M_2}{n^{p+\alpha+1}\rho^n}.$$

This degree of convergence is greater (by the factor $1/n$) than we should have any reason to expect from our general Theorem 12.1.[33] In this sense it is true that analyticity of the function $f(z)$ on C_ρ, except for an algebraic singularity of the kind indicated in (12.20), has the same effect on degree of convergence on C as does a Lipschitz condition on $f^{(p+1)}(z)$ of order α throughout the extent of C_ρ. But it is to be remembered also that the sum of a function of type (12.20) and a function analytic interior to C_ρ, continuous in the corresponding closed region, whose $(p + 1)$-th derivative on C_ρ satisfies a Lipschitz condition of order α, also exhibits degree of convergence for the Tchebycheff polynomial at least as great as that indicated by (12.21). In that sense, it is not necessary to have *analyticity* of $f(z)$ on C_ρ except for the algebraic singularity to ensure (12.21); the Lipschitz condition of order α on $f^{(p+1)}(z)$ is sufficient.

BIBLIOGRAPHY

S. BERNSTEIN.

1. *Sur l'ordre de la meilleure approximation des fonctions continues par des polynomes de degré donné*, Mémoires publiés par l'Académie royale de Belgique, vol. 4(1912), 2e Série, no. 56, pp. 1–105.
2. *Leçons sur les Propriétés Extrémales* ···, Paris, 1926.

JOHN CURTISS.

1. *Interpolation in regularly distributed points*, Transactions of the American Mathematical Society, vol. 38(1935), pp. 458–473.
2. *A note on the degree of polynomial approximation*, Bulletin of the American Mathematical Society, vol. 42(1936), pp. 873–878.

G. DARBOUX.

1. *Sur les développements en série des fonctions d'une seule variable*, Journal de Mathématiques Pures et Appliquées, (3), vol. 2 (1876), pp. 291–312.

P. DIENES.

1. *Sur les Singularités des Fonctions Analytiques*, Paris, 1913.

[33] We have here established (12.21) merely for C a circle, but use of Faber's polynomials rather than Taylor's series extends the result to the case that C is an arbitrary analytic Jordan curve.

G. Faber.

1. *Über polynomische Entwickelungen*, Mathematische Annalen, vol. 57(1903), pp. 389–408.
2. *Über Tchebyscheffsche Polynome*, Journal für die reine und angewandte Mathematik (Crelle), vol. 150(1920), pp. 79–106.

L. Fejér.

1. *Interpolation und konforme Abbildung*, Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse, vol. 30(1918), pp. 319–331.

J. Hadamard.

1. *La série de Taylor*, Scientia, 1901.

G. H. Hardy and J. E. Littlewood.

1. *Some properties of fractional integrals*, II, Mathematische Zeitschrift, vol. 34(1932), pp. 403–439.

Dunham Jackson.

1. *The Theory of Approximation*, American Mathematical Society Colloquium Publications, vol. XI, New York, 1930.
2. *On certain problems of approximation in the complex domain*, Bulletin of the American Mathematical Society, vol. 36(1930), pp. 851–857.
3. *On the application of Markoff's theorem to problems of approximation in the complex domain*, ibid., vol. 37(1931), pp. 883–890.

M. Keldysch and M. Lavrentieff.

1. *Sur la représentation conforme des domaines limités par des courbes rectifiables*, Annales de l'Ecole Normale, vol. 54(1937), pp. 1–38.

O. D. Kellogg.

1. *Harmonic functions and Green's integral*, Transactions of the American Mathematical Society, vol. 13(1912), pp. 109–132.

C. J. de la Vallée Poussin.

1. *Leçons sur l'Approximation des Fonctions d'une Variable Réelle*, Paris, 1919.

E. Lindelöf.

1. *Sur un principe général de l'analyse et ses applications à la théorie de la représentation conforme*, Acta Societatis Scientiarum Fennicae, vol. 46(1915), pp. 1–35.

S. Mandelbrojt.

1. *Modern Researches on the Singularities of Functions Defined by Taylor's Series*, Rice Institute Pamphlet, vol. 14, 1927, pp. 223–352.
2. *La Singularité des Fonctions Analytiques Représentées par une Série de Taylor*, Mémorial des Sciences Mathématiques, vol. 54, 1932.

P. Montel.

1. *Séries de Polynomes*, Paris, 1910.

E. Phragmén and E. Lindelöf.

1. *Sur une extension d'un principe classique de l'analyse*, Acta Mathematica, vol. 31(1908), pp. 381–406.

G. Pólya and G. Szegö.

1. *Aufgaben und Lehrsätze aus der Analysis*, Berlin, 1925.

J. Privaloff.

1. *Sur les fonctions conjugées*, Bulletin de la Société Mathématique de France, vol. 44(1916), pp. 100–103.

S. Saks.

1. *Theory of the Integral*, Warsaw-Lwow, 1937.

W. E. Sewell.

1. *Degree of approximation by polynomials to continuous functions*, Bulletin of the American Mathematical Society, vol. 41(1935), pp. 111–117.

2. *On the modulus of the derivative of a polynomial*, ibid., vol. 42(1936), pp. 699–701.
3. *Generalized derivatives and approximation by polynomials*, Transactions of the American Mathematical Society, vol. 41(1937), pp. 84–123.
4. *Integral approximation and continuity*, Tôhoku Mathematical Journal, vol. 46(1939), pp. 75–78.

J. SHOHAT.
1. *On interpolation*, Annals of Mathematics, (2), vol. 34(1933), pp. 130–146.

V. J. SMIRNOFF.
1. *Sur la théorie des polynomes orthogonaux à une variable complexe*, Journal de la Société Physico-mathématique de Leningrad, vol. 2(1928), pp. 155–178.
2. *Sur formules de Cauchy et de Green et quelques problèmes qui s'y rattachent*, Bulletin de l'Académie des Sciences de l'URSS, vol. 7(1932), pp. 337–371.

G. SZEGÖ.
1. *Orthogonal Polynomials*, American Mathematical Society Colloquium Publications, vol. XXIII, New York, 1939.

J. L. WALSH.
1. *Interpolation and Approximation by Rational Functions in the Complex Domain*, American Mathematical Society Colloquium Publications, vol. XX, New York, 1935.

J. L. WALSH AND J. H. CURTISS.
1. *The Jacobi interpolation series on the lemniscate of convergence*, Bulletin of the American Mathematical Society, vol. 41(1935), p. 25, abstract no. 28.

J. L. WALSH AND W. E. SEWELL.
1. *Note on degree of approximation to an integral by Riemann sums*, American Mathematical Monthly, vol. 44(1937), pp. 155–160.
2. *Note on the relation between continuity and degree of polynomial approximation in the complex domain*, Bulletin of the American Mathematical Society, vol. 43(1937), pp. 557–563.
3. *Note on degree of trigonometric and polynomial approximation to an analytic function*, Bulletin of the American Mathematical Society, vol. 44(1938), pp. 865–873.
4. *Note on degree of trigonometric and polynomial approximation to an analytic function, in the sense of least p-th powers*, Bulletin of the American Mathematical Society, vol. 46(1940), pp. 312–319.

STEFAN WARSCHAWSKI.
1. *Bemerkung zu meiner Arbeit: Über das Randverhalten der Ableitung der Abbildungsfunktion bei konformer Abbildung*, Mathematische Zeitschrift, vol. 38(1934), pp. 669–683.

HARVARD UNIVERSITY AND GEORGIA SCHOOL OF TECHNOLOGY.

INTEGRATION IN ABSTRACT SPACE

By R. L. Jeffery

1. **Introduction.** The problem of integration in abstract space has been approached from many angles.[1] The most interesting, from the standpoint of simplicity and usefulness, is that of Bochner. He considers functions $f(x)$ on a range of measurable sets to a complete normed vector space X. The function $f(x)$ is measurable if it is the limit of a sequence of finite-valued functions, and is integrable if it is measurable and if the real function $\|f(x)\|$ is summable. There are some simple functions to which Bochner's theory does not assign an integral, one of which is:

Let X be the space of bounded functions $x(t)$ on $0 \leqq t \leqq 1$. Let $f(x) = x(t)$, where $x(t) = 0$, $0 \leqq t < x$, $x(t) = 1$, $x \leqq t \leqq 1$, and let $\|f(x)\|$ be the least upper bound of $|x(t)|$, $0 \leqq t \leqq 1$.

The difficulty here is that $f(x)$ is not the limit of a sequence of finite-valued functions. The theory set forth by Birkhoff has all the generality that can reasonably be hoped for. It is, however, somewhat removed from the simplicity that characterizes the work of Bochner and the classical theories for real and complex variables. The present paper started in an attempt to formulate Bochner's definition of a measurable function in terms of the behavior of the function apart from its relation to any sequence. The outcome is a theory of integration which is equivalent to that of Birkhoff, and consequently includes that of Bochner, while the developments are more in the spirit of the classical theories for real and complex variables.

2. **Definition of integrability.** Let X be a complete normed vector space. Let E be any set of elements x on which a measure function has been defined. For the sake of definiteness we take E to be a bounded Lebesgue measurable set in Euclidian space, and x the points of E. Let $f(x)$ be a function defined on E to X. We first define integrability for bounded functions.

Let $f(x)$ be bounded on E. If there exists a sequence of measurable sets

Received March 11, 1940.

[1] L. M. Graves, *Riemann integration and Taylor's theorem in general analysis*, Transactions of the American Mathematical Society, vol. 29(1927), pp. 163–172.

T. H. Hildebrandt, *Lebesgue integration in general analysis*, Bulletin of the American Mathematical Society, vol. 33(1927), p. 646.

S. Bochner, *Integration von Funktionen, deren Wert die Elemente eines Vektorraumes sind*, Fundamenta Mathematicae, vol. 20(1933), pp. 262–276.

N. Dunford, *Integration in general analysis*, Transactions of the American Mathematical Society, vol. 37(1935), pp. 441–453.

Garrett Birkhoff, *Integration of functions with values in a Banach space*, Transactions of the American Mathematical Society, vol. 38(1935), pp. 357–378.

$\Pi^n(E) = e_1^n, e_2^n, \cdots$ with $e_i^n e_j^n = 0, i \neq j, \sum_i me_i^n = mE, me_i^n \to 0$ as $n \to \infty$, and for which

$$\lim_{n\to\infty} f(\xi_i^n)me_i^n \tag{1}$$

exists, where ξ_i^n is any point on e_i^n, then this limit $T = T(f, E)$ is the integral of $f(x)$ over E.

The question at once arises as to the uniqueness of T; for this definition has little interest unless it can be shown that T is the same for all choices of $\Pi^n(E)$ for which the limit (1) exists. This independence of T we establish. We first simplify the notation by writing $\Pi^n(E) = \Pi(E) = e_1, e_2, \cdots, \xi_i^n = \xi_i$, and $\lim \sum_i f(\xi_i^n)me_i^n = \lim \sum f(\xi_i)me_i$, and prove

THEOREM I. *If for $\Pi(E)$ $\lim \sum f(\xi_i)me_i = T$, and for $\Pi'(E)$ $\lim \sum f(\xi_i')m(e_i') = T'$, then $T' = T$.*

We shall need the fundamental

LEMMA I. *Let $B_1, B_2, \cdots, B_n$ be a finite number of bounded sets in X, $\beta_1, \beta_2, \cdots, \beta_n$ elements of the respective sets $B_1, B_2, \cdots, B_n$. Let B be the locus of all vector sums of the form*

$$\beta = \beta_1 + \beta_2 + \cdots + \beta_n,$$

and $\rho(B)$ the least upper bound of $\| \beta - \beta' \|$ for β and β' any two elements of B. Let a_{ij} $(i = 1, 2, \cdots, n; j = 1, 2, \cdots, k_i)$ be sets of positive numbers such that $\sum_i a_{ij} = A$. If β_{ij} $(j = 1, 2, \cdots, k_i)$ are elements of B_i, and β any element of B, then

$$\left\| \sum_{i=1}^{n} \sum_{j=1}^{k_i} a_{ij}\beta_{ij} - A\beta \right\| \leqq A\rho(B).$$

For each value of i let the numbers a_{ij} represent the lengths of intervals placed end to end from left to right in the order $a_{i1}, a_{i2}, \cdots, a_{ik_i}$ over the linear interval $0 \leqq t \leqq A$. For all values of i and j project the end-points of the intervals a_{ij} onto the interval $0 \leqq t \leqq A$. These projections form on $0 \leqq t \leqq A$ a finite set of intervals, $d_1, d_2, \cdots$, and any d_k represents the common part of intervals $a_{ij_1}, a_{ij_2}, \cdots, a_{ij_n}$ for some set of values of $j_1, j_2, \cdots, j_n, j_i$ on the range $1, 2, \cdots, k_i$. Accordingly we set $d_k = d_{j_1\cdots j_n}$. Thus $d_{j_1\cdots j_n}$ is defined for all combinations of values of $j_1 j_2 \cdots j_n$ for which the intervals $a_{ij_1}, a_{ij_2}, \cdots, a_{ij_n}$ have a common part with length > 0. For other combinations of $j_1, j_2, \cdots, j_n$ $(j_i = 1, 2, \cdots, k_i)$ let $d_{j_1\cdots j_n} = 0$. Then $\sum_{j_1\cdots j_n} d_{j_1\cdots j_n} = A$, and

$$\sum_i \sum_{j=1}^{k_i} a_{ij}\beta_{ij} = \sum_{j_1\cdots j_n} d_{j_1\cdots j_n}(\beta_{1j_1} + \beta_{2j_2} + \cdots + \beta_{nj_n}).$$

Since $\beta_{1j_1} + \cdots + \beta_{nj_n}$ is a point $\beta_{j_1 \cdots j_n}$ of B, and $\sum_{j_1 \cdots j_n} d_{j_1 \cdots j_n} = A$, we can write

$$\left\| \sum_{i=1}^{n} \sum_{j=1}^{k_i} a_{ij}\beta_{ij} - A\beta \right\| = \left\| \sum_{j_1 \cdots j_n} d_{j_1 \cdots j_n} (\beta_{j_1 \cdots j_n} - \beta) \right\|$$
$$\leqq \sum_{j_1 \cdots j_n} d_{j_1 \cdots j_n} \| \beta_{j_1 \cdots j_n} - \beta \|$$
$$\leqq A\rho(B).$$

This is the theorem.

COROLLARY I. *Let $e_1, e_2, \cdots, e_n$ be any n disjunct measurable sets on the measurable set E, ξ_i any point on e_i, and $s = \sum f(\xi_i)me_i$, where f is bounded on E. Let $e_{i1}, e_{i2}, \cdots, e_{ik_i}$ be a subdivision of e_i into disjunct measurable sets, ξ_{ij} any point on e_{ij}. Then*

$$\left\| \sum_{i} \sum_{j=1}^{k_i} f(\xi_{ij})\, me_{ij} - s \right\| \leqq \max \left\| \sum \{f(\xi_i) - f(\bar{\xi}_i)\} me_i \right\|,$$

for ξ_i and $\bar{\xi}_i$ any two points on e_i.

Let $f(\xi)me_i = B_i$, where ξ ranges over e_i. Set $me_{ij} = a_{ij}me_i$. Then $\sum_j a_{ij} = 1$, $f(\xi_{ij})me_i$ is a point β_{ij} in B_i and s is a point in B, where B is the locus of all sums of the form $\sum f(\xi_i)me_i$. Hence

$$\sum_{i=1}^{n} \sum_{j=1}^{k_i} f(\xi_{ij})me_{ij} = \sum_{i=1}^{n} \sum_{j=1}^{k_i} a_{ij}f(\xi_{ij})me_i = \sum_{i=1}^{n} \sum_{j=1}^{k_i} a_{ij}\beta_{ij},$$

and since $\sum_j a_{ij} = 1$, Lemma I applies directly to give the corollary.

We now return to the proof of Theorem I. Since f is bounded, there exist sets $e_1, e_2, \cdots$ of $\Pi(E)$, and sets $e_1', e_2', \cdots$ of $\Pi'(E)$, and positive integers n and l such that if we delete from the sets $e_1 + e_2 + \cdots + e_n$, $e_1' + e_2' + \cdots + e_l'$ all the points which are not common to both sets, we have

$$(1) \qquad \left\| \sum_{i=1}^{n} f(\xi_i)me_i - T \right\| < \epsilon, \qquad \left\| \sum_{i=1}^{l} f(\xi_i')me_i' - T' \right\| < \epsilon,$$

where e_i and e_i' now represent the deleted sets, and ξ_i and ξ_i' are any points of e_i and e_i' respectively. From (1) we get

$$(2) \quad \left\| \sum_{i=1}^{n} \{f(\xi_i) - f(\bar{\xi}_i)\} me_i \right\| < 2\epsilon, \qquad \left\| \sum_{i=1}^{l} \{f(\xi_i') - f(\bar{\xi}_i')\} me_i' \right\| < 2\epsilon,$$

where ξ_i, $\bar{\xi}_i$ are any two points on e_i, ξ_i', $\bar{\xi}_i'$ any two points on e_i'. Let $e_{ij} = e_ie_j'$, and ξ_{ij} be any point on e_{ij}. Then by Corollary I, Lemma I, and (2) we have

$$\left\| \sum_{i=1}^{n} \sum_{j=1}^{l} f(\xi_{ij}) m e_{ij} - \sum_{i=1}^{n} f(\xi_i) m e_i \right\| < 2\epsilon,$$

$$\left\| \sum_{i=1}^{n} \sum_{j=1}^{l} f(\xi_{ij}) m e_{ij} - \sum_{i=1}^{l} f(\xi_i') m e_i' \right\| < 2\epsilon.$$

These give

$$\left\| \sum_{i=1}^{n} f(\xi_i) m e_i - \sum_{i=1}^{l} f(\xi_i') m e_i' \right\| < 4\epsilon,$$

and this, with (1), gives $\| T - T' \| < 6\epsilon$. Since ϵ is arbitrary, it follows that $T' = T$, and Theorem I is proved.

3. Extension to unbounded functions.

Let $f(x)$ be finite almost everywhere on E, and let g be the measurable subsets of E over which f is bounded. If for every g $T(f, g)$ exists and there is a point T in X such that

$$\lim_{mg \to mE} T(f, g) = T,$$

then f is integrable over E to the value $T = T(f, E)$.

4. Properties of $T(f, E)$.

THEOREM II. *If E_1 and E_2 are disjunct measurable sets on X, and f is such that $T(f, E_1)$ and $T(f, E_2)$ both exist, then $T(f, E_1 + E_2)$ exists and is equal to $T(f, E_1) + T(f, E_2)$.*

To prove this theorem let g be any measurable set on $E = E_1 + E_2$ over which f is bounded. Write $g = g_1 + g_2$, where $g_1 = gE_1$, $g_2 = gE_2$. Then any sequences $\Pi(g_1)$ and $\Pi(g_2)$ combine to form a sequence $\Pi(g)$, and from this it follows at once that $T(f, g)$ exists and is equal to $T(f, g_1) + T(f, g_2)$. Hence

$$\lim_{mg \to mE} T(f, g) = \lim_{mg_1 \to mE_1} T(f, g_1) + \lim_{mg_2 \to mE_2} T(f, g_2),$$

and from this it follows that $T(f, E)$ exists and $T(f, E) = T(f, E_1) + T(f, E_2)$.

THEOREM III. *Let X be a complete normed vector space, E any bounded measurable set, $f(x)$ a function on E to X and integrable over E. Then f is integrable over every measurable subset e of E.*

Let $g_1, g_2, \cdots$ be a sequence of measurable sets $g_n \subset e$, f bounded on g_n, and $mg_n \to me$, and let $c_1, c_2, \cdots$ be a sequence of measurable sets with $c_n \subset ce$, the complement of e on E, f bounded on c_n and $mc_n \to mce$. Since f is integrable over E, it is integrable over c_n and g_n. Then by Theorem II, f is integrable over $c_n + g_{n'}$ for all n and n', and

$$\int_{c_n + g_{n'}} f\, dx = \int_{c_n} f\, dx + \int_{g_{n'}} f\, dx.$$

By the definition of integrability of f over E there exists N such that

$$\left\| \int_{e_N+g_n} f\,dx - \int_E f\,dx \right\| < \epsilon, \qquad n > N.$$

It then follows that

$$\left\| \int_{g_n} f\,dx - \int_{g_{n'}} f\,dx \right\| < 2\epsilon, \qquad n, n' > N.$$

Consequently $T(f, g_n)$ is a Cauchy sequence in X, and since X is complete, has a limit point in X. It is now necessary to show that this limit is independent of the sequence g_n. Let $g_1', g_2', \cdots$ be another such sequence on e. There then exists N such that

$$\left\| \int_{e_N+g_n} f\,dx - \int_E f\,dx \right\| < \epsilon, \qquad \left\| \int_{e_N+g_n'} f\,dx - \int_E f\,dx \right\| < \epsilon, \qquad n > N$$

Hence

$$\left\| \int_{g_n} f\,dx - \int_{g_n'} f\,dx \right\| < 2\epsilon, \qquad n > N,$$

and this shows that the limit is the same for either sequence. We now can conclude that if g is any measurable set on e over which f is bounded, then $T(f, g)$ tends to a limit as $mg \to me$. Consequently f is integrable over e.

THEOREM IV. *Let f be integrable over E and let e be an element of the measurable subsets of E. Then*

$$\lim_{me\to 0} \int_e f\,dx = 0.$$

By Theorem III f is integrable over both e and ce. If ϵ is given, it follows from the definition of integrability of f over E that if e is a measurable set on E with me sufficiently small, and g is a measurable set on ce such that f is bounded over g and mg is sufficiently close to mce, then

$$\left\| \int_E f\,dx - \int_g f\,dx \right\| < \epsilon, \qquad \left\| \int_{ce} f\,dx - \int_g f\,dx \right\| < \epsilon.$$

These relations give

$$\left\| \int_E f\,dx - \int_{ce} f\,dx \right\| < 2\epsilon. \tag{1}$$

By Theorem II

$$\int_e f\,dx = \int_E f\,dx - \int_{ce} f\,dx,$$

and this with (1) and the fact that ϵ is arbitrary gives the theorem.

THEOREM V. *If E is any measurable set over which f is integrable, and e the measurable subsets of E, then $T(f, e)$ is a completely additive set function over E.*

Let $e_1, e_2, \cdots$ be a sequence of disjunct measurable sets on E, $\sum e_i = e$, and $e_1 + \cdots + e_n = E_n$. By Theorem III, f is integrable over e and E_n, and by Theorem II,

$$\sum_{i=1}^{n} \int_{e_i} f\,dx = \int_{E_n} f\,dx, \qquad \int_{E_n} f\,dx + \int_{e-E_n} f\,dx = \int_e f\,dx = \int_{\Sigma e_i} f\,dx.$$

But $m(e - E_n) \to 0$. Theorem IV and the foregoing equalities now give the desired result.

5. **Integrable functions.** In this section we study classes of integrable functions. There is first defined a class of *restricted* functions, and on this is based a class of *measurable* functions.

The bounded function $f(x)$ on E to X is restricted if there exists a positive real number R such that to any $\epsilon > 0$ there corresponds a sequence $R(E)$ of disjunct measurable sets $e_1, e_2, \cdots$ with $me_i < \epsilon$, $\sum me_i = mE$, and

$$\| \sum_k \{f(\xi_{i_k}) - f(\bar{\xi}_{i_k})\} \| < R,$$

where ξ_{i_k} and $\bar{\xi}_{i_k}$ are any two points on e_{i_k}.

LEMMA II. *If $f(x)$ is restricted on E, then for any sequence $R(E)$*

$$\| \sum \{f(\xi_i) - f(\bar{\xi}_i) me_i \| < 2R \text{ (maximum of } me_i).$$

In considering the sum

$$S_n = \sum_{i=1}^{n} \{f(\xi_i) - f(\bar{\xi}_i)\} me_i$$

there is no loss of generality in assuming that the quantities $me_1, me_2, \cdots, me_n$ are in decreasing order of magnitude. We write

$$f(\xi_i) - f(\bar{\xi}_i) = \varphi_i, \qquad \sum_{i=p}^{n} \varphi_i = A_p.$$

Then $\| A_p \| < R$, and

$$\begin{aligned} S_n &= \varphi_n me_n + \varphi_{n-1} me_{n-1} + \cdots + \varphi_1 me_1 \\ &= A_n me_n + (A_{n-1} - A_n) me_{n-1} + \cdots + (A_1 - A_2) me_1 \\ &= - A_n(me_{n-1} - me_n) - A_{n-1}(me_{n-2} - me_{n-1}) - \cdots - A_2(me_1 - me_2) \\ &\qquad + A_1 me_1 < R(2me_1 - me_n) < 2Rme_1. \end{aligned}$$

Since this relation holds for all values of n, the lemma follows.

THEOREM VI. *If the bounded function $f(x)$ is restricted, it is integrable.*

By Lemma II there exists $R^n(E) = e_1^n, e_2^n, \cdots$ such that if

$$S_n = \sum_i \{f(\xi_i^n) - f(\bar{\xi}_i^n)\} me_i^n ,$$

where ξ_i^n and $\bar{\xi}_i^n$ are any two points on e_i^n, then $\| S_n \| \to 0$ as $n \to \infty$. Since $f(x)$ is bounded and X is complete, $\sum_i f(\xi_i^n) me_i^n$ converges. Set

$$\sum_i f(\xi_i^n) me_i^n = T_n , \qquad \sum_i f(\xi_i^{n\prime}) me_i^{n\prime} = T_{n'} .$$

Since $\| S_n \| \to 0$ the methods of proof used in Lemma I and Theorem I can be used to show that for n and n' sufficiently great $\| T_n - T_{n'} \|$ is arbitrarily small. Consequently T_n forms a Cauchy sequence in X and T_n tends to a limit T. But $R^n(E)$ is a sequence $\Pi(E)$ of the definition of integrability, and it follows that $f(x)$ is integrable.

Corollary. *The function $f(x)$ of the example given in the introduction is integrable.*

For if $e_1, \cdots, e_n$ is a subdivision of $0 \leqq x \leqq 1$ into n equal parts, it is easily verified that for every choice of i_k and ξ_{i_k}, $\bar{\xi}_{i_k}$ on e_{i_k},

$$\| \sum_k \{f(\xi_{i_k}) - f(\bar{\xi}_{i_k})\} \| = 1.$$

This shows that $f(x)$ is restricted. The integrability of f then follows from the theorem.

Let $f(x)$ be a function defined on the measurable set E to the complete normed vector space X. The function $f(x)$ is measurable if it is the limit of a sequence of restricted functions.

It follows from this definition that if $f(x)$ is measurable on E, it is measurable on every measurable subset e of E. For it is obvious that if a function is restricted on E, it is restricted on every measurable subset of E.

Theorem VII. *If $f(x)$ is bounded and measurable, it is integrable.*

Let $f_n(x)$ be the sequence of restricted functions whose limit is $f(x)$ and R_n the number associated with $f_n(x)$ in the definition of restricted functions. Let $\epsilon_1, \epsilon_2, \cdots$ be a sequence of positive numbers tending to zero. There then exists a sequence $e_1^n, e_2^n, \cdots$ of disjunct measurable sets on E with $me_i^n < \epsilon_n/2R_n$, $\sum me_i^n = mE$, and with

$$\text{(1)} \qquad \{f_n(\xi_i^n) - f_n(\bar{\xi}_i^n)\} me_i^n < 2R_n \text{ (maximum of } me_i^n) < \epsilon_n ,$$

where ξ_i^n and $\bar{\xi}_i^n$ are any two points on e_i^n. Since $f_n(x) \to f(x)$, for a given η there exists a set $E_\eta \subset E$ with $mE - mE_\eta < \eta$, and such that $f_n \to f(x)$ uniformly on E_η. Let $E_i^n = e_i^n E_\eta$. Then (1) holds with the sets E_i^n replacing e_i^n and with ξ_i^n and $\bar{\xi}_i^n$ any two points on E_i^n. We now have

$$\| \sum_i \{f(\xi_i^n) - f(\bar{\xi}_i^n)\} \| \leqq \sum_i \{\| f(\xi_i^n) - f_n(\xi_i^n) \| + \| f_n(\bar{\xi}_i^n) - f(\bar{\xi}_i^n) \| + \| f_n(\xi_i^n) - f_n(\bar{\xi}_i^n) \| \}.$$

Let λ_n be the maximum of $\| f(x) - f_n(x) \|$ for x on E_η. Then since on E_η $f_n(x) \to f(x)$ uniformly, it follows that $\lambda_n \to 0$ as $n \to \infty$. In the foregoing inequality the points ξ_i^n, $\bar{\xi}_i^n$ are on E_η for all values of i and n. Taking (1) into consideration, we are able to write

$$\| \sum_i \{f(\xi_i^n) - f(\bar{\xi}_i^n)\} mE_i^n \| < 2\lambda_n mE_\eta + \epsilon_n,$$

where λ_n and ϵ_n both tend to zero as $n \to \infty$. From this point on, the proof goes as in Theorem VI to show that the sequence $E_1^n, E_2^n, \cdots$ is a sequence $\pi(E_\eta)$ and consequently $f(x)$ is integrable over E_η. Since $mE_\eta \to mE$ and $f(x)$ is bounded on E, it is not difficult to show that $f(x)$ is integrable over E.

Corollary I. *If $f(x)$ is measurable on E, and g is a measurable set with $g \subset E$ and $f(x)$ bounded over g, then $f(x)$ is integrable over g.*

Following the definition of measurability of $f(x)$ it was noted that if $f(x)$ is measurable on E it is measurable on every measurable subset e of E. With g measurable and f bounded on g the corollary follows immediately from the theorem.

Corollary II. *If $f(x)$ is measurable over E and $\| f(x) \|$ is summable, then $f(x)$ is integrable over E.*

By Corollary I, $f(x)$ is integrable over the measurable sets $g \subset E$ on which $f(x)$ is bounded. The fact that $\| f(x) \|$ is summable easily leads to the conclusion that

$$\lim_{mg \to mE} \int_g f \, dx$$

exists. Then, by definition, f is integrable over E.

Corollary III. *If $f(x)$ is the limit of a sequence of finite-valued functions, then $f(x)$ is measurable.*

For if $f_n(x)$ is a finite-valued function on E, then for a given ϵ the set E can be divided into a finite number of sets $e_1, e_2, \cdots, e_k$ with $me_i < \epsilon$, and such that $f(x)$ is constant over e_i. Hence if ξ_i, $\bar{\xi}_i$ are any two points on e_i

$$\| f_n(\xi_i) - f_n(\bar{\xi}_i) \| = 0,$$

and we conclude that $f_n(x)$ is restricted.

Corollary IV. *If $f(x)$ is integrable in Bochner's sense, it is integrable.*

For if $f(x)$ is integrable in Bochner's sense, it is the limit of a sequence of finite-valued functions and $\| f(x) \|$ is summable. Corollary IV then follows from Corollaries II and III. That the converse does not hold follows from the corollary to Theorem VI.

We close this section with some comments on restricted functions and on Bochner's integral. We have not been able to arrive at any idea as to how wide is the class of restricted functions. It was pointed out in the corollary to

Theorem IV that the function defined in the introduction was restricted. If $f(x)$ is a bounded Lebesgue measurable function of the real variable x and $\| f(x) \| = | f(x) |$, then $f(x)$ is restricted. To show this we first let $\alpha = \alpha_1$, $\alpha_2, \cdots$ be the denumerable set of disjunct sets which are such that $\alpha_i \subset E$, $m\alpha_i > 0$, and $f(x)$ is constant on α_i. Each set α_i can then be subdivided into a finite number of sets $\alpha_{i1}, \cdots, \alpha_{ik_i}$ such that $m\alpha_{ij} < \epsilon$. If ξ_{ij} and $\bar{\xi}_{ij}$ are any two points on α_{ij}, then

$$\| \sum \{f(\xi_{ij}) - f(\bar{\xi}_{ij})\} \| = 0. \tag{1}$$

Divide the range of f on $E - \alpha$ by the points $a_0, a_1, \cdots, a_n$ where $a_i - a_{i-1} \to 0$ as $n \to \infty$, and let e_i be the part of $E - \alpha$ for which $a_{i-1} \leqq f < a_i$ ($i = 1, 2, \cdots, n - 1$), e_n the set for which $a_{n-1} \leqq f \leqq a_n$. Then for any choice of ξ_i and $\bar{\xi}_i$ on e_i

$$| \sum f(\xi_i) - f(\bar{\xi}_i) | \leqq M - m, \tag{2}$$

where M is the least upper bound, m the greatest lower bound of $f(x)$ on $E - \alpha$. It will next be shown that $me_i \to 0$ as $a_i - a_{i-1} \to 0$. For if this is not the case, there exists a number $d > 0$ and an infinite set of values of n and i for which $e_i = e_i^n$ is such that $me_i^n > d$. Let e be the set which is such that each point of e is in an infinite number of the sets e_i^n. Then[2] $me \geqq d$. Let x be a point of e. Then x is in an infinite number of sets $e_j = e_{i_j}^{n_j}$ with $me_j > d$, and such that if ξ_j is a point of e_j distinct from x, then

$$| f(x) - f(\xi_j) | \leqq a_{i_j}^{n_j} - a_{i_j-1}^{n_j} = \eta_j,$$

where $\eta_j \to 0$ as $n_j \to \infty$. If e_x is the set which is such that each of its points is in an infinite number of the sets e_j, then $me_x \geqq d$, and for every point ξ of e_x, $f(\xi) = f(x)$. Consequently $f(x)$ is constant on e_x, and since $me_x > 0$, e_x belongs to α. But this is a contradiction and we conclude that $me_i \to 0$ as $a_i - a_{i-1} \to 0$. If (1) and (2) are taken into consideration, it now follows that the sets α_{ij}, and the sets e_i for $a_i - a_{i-1}$ sufficiently small can be combined to form a sequence of disjunct sets $e_1', e_2', \cdots$ with $me_i' < \epsilon$, $\sum me_i' = mE$, and such that if ξ_i' and $\bar{\xi}_i'$ are any two points on e_i' then

$$| \sum f(\xi_i') - f(\bar{\xi}_i') | < M - m,$$

and consequently $f(x)$ is restricted.

If $f(z) = u(x, y) + iv(x, y)$ is a function of the complex variable $z = x + iy$ for which $u(x, y)$ and $v(x, y)$ are bounded and measurable on the measurable set E of the complex plane, then it can be shown in a similar manner that $f(z)$ is restricted. Beyond this we have, so far, been unable to go. E.g., we have not shown that if $f(x)$ is a continuous function on the linear interval $a \leqq x \leqq b$ to the complete normed vector space x, then $f(x)$ is restricted. Such a function

[2] Hobson, *Theory of Functions of a Real Variable*, 3d ed., vol. I, §136.

is measurable in the sense defined above for the reason that it is the limit of a sequence of finite-valued functions.

There is a further question concerning Bochner's integral to which we have given some consideration without arriving at an answer. In the introduction to this paper an example is given of a bounded integrable function which is not integrable in Bochner's sense for the reason that it is not the limit of a sequence of finite-valued functions. Does there exist an integrable function which is not integrable in Bochner's sense for the reason that $\| f(x) \|$ is not summable? We leave a further consideration of this to be included in a study of conditionally convergent integrals with which it seems to be closely associated.

It was also mentioned in the introduction that this paper originated in an attempt to formulate Bochner's definition of measurability of a function without considering the relation of the function to any sequence.

Let $f(x)$ be defined on E to the complete normed vector space X. If for a given $\epsilon > 0$ there exists a sequence of disjunct measurable sets $e_1, e_2, \cdots$ with $me_i < \epsilon$, $e_i \subset E$, $\sum me_i = mE$, and for ξ_i, $\bar{\xi}_i$ any two points on e_i, $\| f(\xi_i) - f(\bar{\xi}_i) \| < \epsilon$, then $f(x)$ is measurable on E.

We shall call this ϵ-measurability to distinguish it from the definition of measurability given above, and state the following results, which are not difficult to prove.

Measurability in the sense of Bochner and ϵ-measurability are equivalent.

Continuous functions are ϵ-measurable.

A necessary and sufficient condition that $f(x)$ be ϵ-measurable is that corresponding to a given $\eta > 0$ there exists a measurable set $E_\eta \subset E$ with $mE - mE_\eta < \eta$, and such that $f(x)$ is ϵ-measurable on E_η.

If $f(x)$ is ϵ-measurable on the linear interval $a \leqq x \leqq b$, then $f'(x)$ is ϵ-measurable, where $f'(x)$ is a function for which

$$\lim_{h \to 0} \left\| \frac{f(x+h) - f(x)}{h} - f'(x) \right\| = 0.$$

5. The equivalence of $T(f, E)$ and Birkhoff's integral $B(f, E)$.

THEOREM VI. *If $B(f, E)$ exists, then $T(f, E)$ exists, and the two are equal.*

If the function $f(x)$ is bounded and $B(f, E)$ exists, it follows from the way $B(f, E)$ is defined that there exists a sequence of subdivisions of E into measurable sets $e_1^n, e_2^n, \cdots$ such that

$$B(f, E) = \lim_{n \to \infty} \sum_i f(\xi_i^n) me_i^n,$$

where ξ_i^n is any point on e_i^n. In this determination of $B(f, E)$ it is not necessary that $me_i^n \to 0$ as $n \to \infty$. Let e_i^n be broken into measurable subsets e_{i1}^n, $e_{i2}^n, \cdots, e_{ik_i}^n$ in such a way that $me_{ij}^n \to 0$ as $n \to \infty$. It then follows from Corollary I, Lemma I of this paper that

$$\lim_{n\to\infty} \sum_i \sum_{j=1}^{k_i} f(\xi_{ij}^n) me_{ij}^n = B(f, E).$$

These sets e_{ij}^n can be arranged in a single sequence $\Pi^n(E) = e_1^n, e_2^n, \cdots$ which constitutes a sequence $\Pi(E)$ of E for which $\lim \sum f(\xi_i) me_i = B(f, E)$. From this and the definition of $T(f, E)$ it follows that $T(f, E)$ exists and is equal to $B(f, E)$.

Coming to the case for which $f(x)$ is unbounded, we note that if $B(f, E)$ exists then $B(f, e)$ exists for every measurable subset $e \subset E$, and $B(f, e)$ is a completely additive set function over[3] E. It then easily follows that as $me \to mE$, $B(f, e) \to B(f, E)$. Hence, if g is a measurable subset of E over which f is bounded, then $B(f, g)$ exists, and from the first part of the theorem it follows that $T(f, g)$ exists and is equal to $B(f, g)$. From this and the foregoing remarks concerning $B(f, E)$, together with the definition of $T(f, E)$ for unbounded functions, it follows that

$$\lim_{mg\to mE} B(f, g) = B(f, E) = \lim_{mg\to mE} T(f, g) = T(f, E).$$

This establishes the theorem.

THEOREM VII. *If $T(f, E)$ exists, then $B(f, E)$ exists, and the two are equal.*

If the function $f(x)$ is bounded, then

$$T(f, E) = \int_E f\,dx = \lim_{n\to\infty} \sum_i f(\xi_i^n) me_i^n, \tag{1}$$

where e_i^n is measurable, $e_i^n \neq e_j^n$, $i \neq j$, $\sum me_i^n = mE$, ξ_i^n is any point on e_i^n. Set $e_0^n = E - \sum_i e_i^n$. Then $me_0^n = 0$, and consequently, for any choice of ξ_0^n on e_0^n, $\|f(\xi_0^n) me_0^n\| = 0$. This and the boundedness of f insures that for this subdivision $e_0^n, e_1^n, \cdots$ of E, $\sum_i f(\xi_i^n) me_i^n$ is for each n unconditionally convergent, and from (1) it follows that

$$\lim_{n\to\infty} \| \sum_i \{f(\xi_i^n) - f(\bar{\xi}_i^n)\} me_i^n \| = 0, \tag{2}$$

for ξ_i^n and $\bar{\xi}_i^n$ any two points on e_i^n. From (2) it follows that there are integral ranges[4] of f of arbitrarily small diameter. Relations (1) and (2) then combine to show that the point of intersection of these integral ranges is $T(f, E)$. Consequently $B(f, E)$ exists[5] and is equal to $T(f, E)$.

If the function $f(x)$ is unbounded, the proof is not so simple. Since f is finite almost everywhere on E, it is easy to establish the existence of a sequence of disjunct measurable sets $\mathcal{E}_1, \mathcal{E}_2, \cdots$ with f bounded on each $\mathcal{E}_i$ and $\sum m\mathcal{E}_i = mE$. Since f is integrable over E, the same holds for each $\mathcal{E}_i$. Then, since f

[3] Birkhoff, loc. cit., p. 367, Theorem 14.
[4] Birkhoff, loc. cit., p. 367, Definition 2.
[5] Birkhoff, loc. cit., p. 367, Definitions 3, 4, and Theorems 12, 13.

is bounded on each $\mathfrak{S}_i$, it follows that corresponding to $\epsilon > 0$ and ϵ_i with $\sum \epsilon_i = \epsilon$ there exists a sequence of disjunct measurable sets $e_{i1}, e_{i2}, \cdots$ on $\mathfrak{S}_i$ with $me_{ij} < \epsilon$, $\sum_j me_{ij} = m\mathfrak{S}_i$, and such that if ξ_{ij} is any point on e_{ij}, then

$$(3) \qquad \left\| \sum_j f(\xi_{ij})me_{ij} - \int_{\mathfrak{S}_i} f\,dx \right\| < \epsilon_i,$$

and consequently

$$(4) \qquad \| \sum_j \{f(\xi_{ij}) - f(\bar{\xi}_{ij})\}\, me_{ij} \| < 2\epsilon_i,$$

for ξ_{ij}, $\bar{\xi}_{ij}$ any two points on e_{ij}. We next show that the series $\sum_{ij} f(\xi_{ij})me_{ij}$ converges unconditionally. With this established it follows from (4) and the definition of integral ranges that there exist integral ranges of f of arbitrarily small diameter, and this in turn gives the existence of $B(f, E)$.

The unconditional convergence of the series $\sum_{ij} f(\xi_{ij})me_{ij}$ follows if for a given ϵ there is established the existence of an integer N such that if $i_k, j_k > N$ then

$$(5) \qquad \| \sum_k f(\xi_{i_kj_k})\, me_{i_kj_k} \| < \epsilon.$$

The existence of this number N we now proceed to show. Let $\eta > 0$ be given. Fix δ such that if e is any measurable set on E with $me < \delta$, then

$$(6) \qquad \left\| \int_e f\,dx \right\| < \eta.$$

Theorem IV permits this. Next fix a positive integer N so that

$$(7) \qquad \sum_{i,j=N+1}^{\infty} me_{ij} < \delta, \qquad \sum_{i=N+1}^{\infty} \epsilon_i < \tfrac{1}{2}\eta.$$

It follows from (4) that for any choice of ξ_{ij_k}, $\bar{\xi}_{ij_k}$ on e_{ij_k},

$$(8) \qquad \| \sum_k \{f(\xi_{ij_k}) - f(\bar{\xi}_{ij_k})\}\, me_{ij_k} \| < 2\epsilon_i.$$

To see this let ξ_{ij_k}, $\bar{\xi}_{ij_k}$ have any values on the sets e_{ij_k} and on the remaining sets of the sequence e_{ij} in (4) let $\xi_{ij} = \bar{\xi}_{ij}$.

Now consider any subsequence $e_{i_kj_k}$ of e_{ij} with $i_k, j_k > N$, and set $e_k = e_{i_kj_k}$. Since e_k belongs to some $\mathfrak{S}_i$, f is bounded on e_k. It then follows from the definition of integrability of bounded functions that there exists a sequence of disjunct measurable sets $e_{k1}, e_{k2}, \cdots$ on e_k with $\sum_l me_{kl} = me_k$ for which

$$(9) \qquad \left\| \sum_l f(\xi_{kl})\, me_{kl} - \int_{e_k} f\,dx \right\| < \eta_k,$$

where ξ_{kl} is any point on e_{kl}, and η_k has been so fixed that $\sum \eta_k < \eta$. If

these sets e_{kl} are formed for all values of k, it follows from Corollary I, Lemma I, as in the proof of Theorem I, that

$$(10) \qquad \left\| \sum_{kl} f(\xi_{kl})\, me_{kl} - \sum_{k} f(\xi_k)\, me_k \right\| \leqq \max \left\| \sum \{f(\xi_k) - f(\bar{\xi}_k)\}\, me_k \right\|$$

for ξ_k, $\bar{\xi}_k$ any two points on e_k. But from (7) and (8) we have, if $i_k > N$,

$$(11) \qquad \left\| \sum \{f(\xi_k) - f(\bar{\xi}_k)\}\, me_k \right\| = \left\| \sum_k f(\xi_{i_k j_k}) - f(\bar{\xi}_{i_k j_k})\, me_{i_k j_k} \right\| < 2 \sum_{N+1}^{\infty} \epsilon_i < \eta.$$

If $i_k, j_k > N$, it follows from (6), (7), and (9) that

$$\left\| \sum_{kl} f(\xi_{kl})\, me_{kl} \right\| < \left\| \sum \int_{e_k} f\, dx + \sum \eta_k \right\| < 2\eta.$$

Relations (10) and (11) then combine to show that

$$\left\| \sum_k f(\xi_{i_k j_k})\, me_{i_k j_k} \right\| = \left\| \sum f(\xi_k)\, me_k \right\| < 3\eta.$$

Since η is arbitrary, the unconditional convergence of $\sum_{ij} f(\xi_{ij}) me_{ij}$ is established.

From the fact that $T(f, e)$ is completely additive over E it follows that

$$\sum \int_{\mathcal{E}_i} f\, dx = \int_E f\, dx,$$

and this, with (1), gives

$$(12) \qquad \left\| \sum_{ij} f(\xi_{ij}) me_{ij} - \int_E f\, dx \right\| < \epsilon.$$

From (4) we have

$$(13) \qquad \left\| \sum_{ij} f(\xi_{ij}) - f(\bar{\xi}_{ij}) me_{ij} \right\| < 2 \sum \epsilon_i < 2\epsilon.$$

It follows from (13) that there are integral ranges of f of arbitrarily small diameter. If a sequence of positive numbers $\epsilon_1, \epsilon_2, \cdots$ is taken with $\epsilon_n \to 0$, and for each ϵ_n the sets e_{ij} of (12) are determined, ϵ_n replacing ϵ in (12), then

$$\lim_{\epsilon_n \to 0} \sum_{ij} f(\xi_{ij}) me_{ij} = \int_E f\, dx.$$

This, with (13), shows that the point of intersection of the integral ranges of f is $T(f, E)$. But by definition the point of intersection of the integral ranges of f is $B(f, E)$. We can now conclude that $B(f, E)$ exists and is equal to $T(f, E)$.

Theorems VI and VII establish the equivalence of Birkhoff's theory and the theory developed in this paper.

ACADIA UNIVERSITY.

UNIVALENT DERIVATIVES OF ENTIRE FUNCTIONS

By R. P. Boas, Jr.

Let $f(z)$ be an entire function, and let $M(r)$ denote the maximum of $|f(z)|$ in $|z| \leqq r$. The object of this note is to establish the existence of a positive number T for which the following theorem is true.

Theorem. *If the entire function $f(z)$ satisfies*

$$\limsup_{r\to\infty} \frac{1}{r} \log M(r) < T, \tag{1}$$

and $f(z)$ is not a polynomial, an infinite number of the derivatives of $f(z)$ are univalent in the unit circle, $|z| \leqq 1$.

It will be shown that a possible value for T is $\log 2$; I do not know whether or not this is the best possible value.

The following corollary is immediately obtainable by a change of variable and an application of the diagonal process.

Corollary. *If $f(z)$ is an entire function, not a polynomial, of order less than one, or of order one and minimum type, then corresponding to any increasing sequence of numbers r_n there is an increasing sequence of integers k_n such that $f^{(k_n)}(z)$ is univalent in $|z| < r_n$ $(n = 1, 2, \cdots)$.*

I show first that if $f(z)$ satisfies (1), with sufficiently small T, and neither $f(z)$ nor any derivative is univalent in the unit circle, then $f(z)$ is a constant. If neither $f(z)$ nor any derivative is univalent, there exist numbers a_n, b_n, such that for $n = 1, 2, \cdots$,

$$\begin{gathered} |a_n| \leqq 1, \; |b_n| \leqq 1, \; a_n \neq b_n, \\ f^{(n-1)}(a_n) = f^{(n-1)}(b_n). \end{gathered} \tag{2}$$

Without loss of generality, we may assume

$$f(0) = 0. \tag{3}$$

Consider the functions $h_n(z)$ defined as follows:

$$h_0(0) = 0,$$

$$z^n[1 + h_n(z)] = z^{n-1}\frac{e^{a_n z} - e^{b_n z}}{a_n - b_n} \qquad (n = 1, 2, \cdots).$$

It is obvious that $h_n(0) = 0$, since

$$1 + h_n(z) \to 1 \qquad (z \to 0).$$

Received February 23, 1940.

Furthermore,

$$h_n(z) = \frac{e^{a_n z} - e^{b_n z}}{(a_n - b_n)z} - 1$$
$$= \sum_{k=2}^{\infty} \frac{(a_n z)^k - (b_n z)^k}{k!(a_n z - b_n z)};$$

hence, since $|a_n| \leqq 1$ and $|b_n| \leqq 1$, $h_n(z)$ is majorized by

$$h(z) = \sum_{k=2}^{\infty} \frac{kz^{k-1}}{k!} = e^z - 1.$$

Hence, for $|z| \leqq r$,

$$|h_n(z)| \leqq h(r) = e^r - 1. \tag{4}$$

We are now in a position to apply a particular case of a theorem of S. Takenaka, which is a sharpened form of an older theorem of S. Pincherle; we state what we need as a lemma.

LEMMA.[1] *If the functions $h_n(z)$ are analytic in $|z| < r$, vanish at $z = 0$, and satisfy*

$$|h_n(z)| \leqq N(s) \qquad (|z| \leqq s < r;\ n = 0, 1, \cdots),$$

then any function $G(z)$, analytic in $|z| < r$, has an expansion of the form

$$G(z) = \sum_{n=0}^{\infty} c_n z^n [1 + h_n(z)], \tag{5}$$

converging uniformly in any circle

$$|z| \leqq R < \sup_{0<s<r} \frac{s}{1 + N(s)}. \tag{6}$$

Our functions $h_n(z)$ satisfy the conditions of the lemma with arbitrary r, and $N(s) = e^s - 1$. In this case

$$\sup_{s>0} \frac{s}{1 + N(s)} = \frac{1}{e}. \tag{7}$$

The lemma, applied to the function $G(z) = e^{zw}$, shows that

$$e^{zw} = \sum_{n=0}^{\infty} c_n(w) z^n [1 + h_n(z)] \qquad (|z| < 1/e), \tag{8}$$

[1] S. Takenaka, *On the expansion of analytic functions in series of analytic functions and its application to the study of the distribution of the zero points of the derivatives of analytic functions*, Nippon Sûgaku-Buturigakkwai Kizi (Proceedings of the Physico-Mathematical Society of Japan), (3), vol. 13(1931), pp. 111–132.

G. S. Ketchum, *On certain generalizations of the Cauchy-Taylor expansion theory*, Transactions of the American Mathematical Society, vol. 40(1936), pp. 208–224; p. 214.

the series being uniformly convergent, for each w, in any circle $|z| \leqq T' < 1/e$. Now by a well-known theorem[2] an entire function $f(w)$ satisfying (1) can be represented in the form

$$f(w) = \int_C e^{zw} F(z)\, dz, \tag{9}$$

where

$$\limsup_{r\to\infty} \frac{1}{r} \log M(r) = t < T;$$

C is a circle $|z| = t'$, $T > t' > t$; and $F(z)$ is analytic on C. We substitute the expansion (8) into (9), and integrate term by term, obtaining

$$\begin{aligned} f(w) &= \sum_{n=0}^{\infty} c_n(w) \int_C z^n[1 + h_n(z)]F(z)\, dz \\ &= c_0(w) \int_C F(z)\, dz + \sum_{n=1}^{\infty} \frac{c_n(w)}{a_n - b_n} \int_C z^{n-1}[e^{a_n z} - e^{b_n z}]F(z)\, dz \\ &= c_0(w)f(0) + \sum_{n=1}^{\infty} \frac{c_n(w)}{a_n - b_n} [f^{(n-1)}(a_n) - f^{(n-1)}(b_n)] \\ &= 0, \end{aligned}$$

since $f(z)$ satisfies (2) and (3). That is, we have shown that if $f(z)$ satisfies (1), with $T = 1/e$, and if neither $f(z)$ nor any derivative is univalent in $|z| \leqq 1$, then $f(z)$ is a constant.

If $f(z)$ has only a finite number of univalent derivatives, let k be the order of the last univalent one. Then the function $f^{(k+1)}(z)$ satisfies (1) and hence, by what has already been proved, is a constant. Consequently $f(z)$ is a polynomial of degree at most $k + 1$.

I have shown in a paper not yet published[3] that, if the functions $h_n(z)$ of the lemma have a common majorant $h(z)$, the expansion of the lemma is valid in any circle in which $|h(z)| < 1$. In establishing (4), we showed that all the $h_n(z)$ are majorized by the function $h(z) = e^z - 1$; we then have $|h(z)| < 1$ for $|z| < \log 2$. Hence we may take $T = \log 2$ in the theorem.

DUKE UNIVERSITY.

[2] G. Pólya, *Untersuchungen über Lücken und Singularitäten von Potenzreihen*, Mathematische Zeitschrift, vol. 29(1929), pp. 549–640; pp. 580 ff.

[3] *Expansions of analytic functions*, to appear in the Transactions of the American Mathematical Society. The result is implicit in my preliminary note, *General expansion theorems*, Proceedings of the National Academy of Sciences, vol. 26(1940), pp. 139–143.

RESTRICTIONS IMPOSED BY CERTAIN FUNCTIONS ON THEIR FOURIER TRANSFORMS

By N. Levinson

1. It is our purpose here to consider the restrictions imposed by the special behavior of a function on its Fourier transform. We shall consider two cases: (a) where the function has special behavior at infinity, and (b) where the function has special behavior at some finite point.

Case (a). The results here began with a suggestion of Wiener that both a function and its Fourier transform cannot be very small at infinity. This suggestion led to a theorem by Hardy,[1] a corollary of which is the fact that *if*

$$f(x) = O(|x|^n e^{-\frac{1}{2}x^2}) \qquad (x \to \pm\infty),$$

and if the Fourier transform of $f(x)$

$$g(u) = o(e^{-\frac{1}{2}u^2}) \qquad (u \to \pm\infty),$$

then $f(x) \equiv 0$.

This result is extended in a theorem of Morgan[2] who shows that *if*

$$f(x) = O(e^{-A|x|^p}) \qquad (x \to \pm\infty;\ p \geqq 2),$$

and its transform

$$g(u) = O(\exp\{-[A' + \epsilon]\,|u|^{p'}\}) \qquad (u \to \pm\infty),$$

where $\epsilon > 0$,

$$\frac{1}{p} + \frac{1}{p'} = 1,$$

and

$$A' = \frac{1}{p'(Ap)^{p'-1}} \sin\frac{\pi}{2(p-1)},$$

then $f(x) \equiv 0$.

The results we shall consider here, while obviously related to the above results, will differ from them in that first we shall restrict the behavior of $f(x)$ and $g(u)$ on only one side at infinity, for example, only as $x \to +\infty$ and $u \to +\infty$.

Received April 13, 1940.

[1] G. H. Hardy, *A theorem concerning Fourier transforms*, Journal London Math. Soc., vol. 8(1933), pp. 227–231.

[2] G. W. Morgan, *A note on Fourier transforms*, Journal London Math. Soc., vol. 9(1934), pp. 187–192.

Secondly, we shall consider somewhat more general behavior at infinity than an exponential of algebraic power. That is, here we shall consider the case where

$$f(x) = O(e^{-h(x)}) \qquad (x \to +\infty)$$

and

$$g(u) = O(e^{-t(u)}) \qquad (u \to +\infty),$$

$g(u)$ being the Fourier transform of $f(x)$, imply that $f(x)$ is zero almost everywhere, where $h(x)$ can be $\frac{1}{2}x^2$ as in the Hardy theorem or e^x or e^{e^x}. In this way the particular case when $h(x) = \infty$ will give the following theorem (which is an immediate consequence of a theorem of Carleman, and was given in more precise form than below by Wiener and Paley).

If

$$f(x) = 0 \qquad (x \to +\infty)$$

and its Fourier transform

$$g(u) = O(e^{-\theta(u)}) \qquad (u \to +\infty),$$

where $\theta(u)$ is increasing and

$$\int_1^\infty \frac{\theta(u)}{u^2}\, du = \infty, \tag{1.1}$$

then $f(x)$ is zero almost everywhere.

[Throughout this paper it will be obvious that *in any formula in the statement of the theorems $+\infty$ can be replaced by $-\infty$.*]

Case (b). Here we consider the case where $f(x)$ becomes small for some finite value of x. The extreme case where $f(x)$ vanishes over an interval has been given in the following theorem:[3] *If $f(x) \in L(-\infty, \infty)$ and if $f(x)$ vanishes over any interval and its Fourier transform*

$$g(u) = O(e^{-\theta(u)}) \qquad (u \to +\infty),$$

where $\theta(u)$ is increasing and satisfies (1.1), *then $f(x)$ vanishes almost everywhere.*

An example of the type of result which will be covered here is that if

$$f(x) = O(\exp\{-\exp[\log x^{-1}]^2\}) \qquad (x \to +0),$$

and

$$g(u) = O(\exp\{-\theta(u) - \pi u\,(\log u)^{-\frac{1}{2}}\}) \qquad (u \to +\infty),$$

then $f(x)$ vanishes almost everywhere.

The extreme case where $f(x)$ vanishes over an interval will appear as a limiting case of a general theorem which will be given in §3.

[3] N. Levinson, *On a class of non-vanishing functions*, Proc. London Math. Soc., vol. 41(1936), p. 393, Theorem 11.

2. THEOREM I. *Let $f(x) \epsilon L(-\infty, \infty)$ and let*

$$f(x) = O(e^{-h(x)}) \qquad (x \to +\infty), \tag{2.01}$$

where $h(x)$ is a positive increasing and twice differentiable function and

$$h''(x) \geqq c > 0 \tag{2.02}$$

for some c. Let the inverse function of $h'(x)$ be $H(x)$. Let $g(u)$, the Fourier transform of $f(x)$, satisfy

$$g(u) = O(e^{-t(u)}) \qquad (u \to +\infty). \tag{2.03}$$

If as $x \to \infty$

$$\int_1^x \left(\frac{1}{y^2} - \frac{1}{R^2}\right) t(y)\, dy - \pi H(x) \to \infty, \tag{2.04}$$

then $f(x)$ vanishes almost everywhere.

Theorem I is an immediate consequence of the following theorem.

THEOREM II. *In Theorem* I, (2.03) *and* (2.04) *can be replaced by*

$$\varliminf_{x\to\infty} \left\{\left(\int_{-x}^{-1} + \int_1^x\right)\left(\frac{1}{y^2} - \frac{1}{x^2}\right) \log |g(y)|\, dy + \pi H(x)\right\} = -\infty. \tag{2.05}$$

Actually (2.04) is a very precise result. We shall show in Theorem III that the constant π cannot be improved for general $h(x)$. However, in case $h(x)$ is of algebraic growth, the constant is not best possible. Thus if $h(x) = \frac{1}{2}x^2$, $H(x) = x$ and (2.04) requires that $t(u) > \frac{3}{2}\pi u^2$, whereas from the theorem of Hardy quoted above we should expect $t(u) > \frac{1}{2}u^2$. We shall show that this is the case by giving a sharper statement of (2.04) when $h(x) = Ax^p$. This will be done in §4.

We now prove Theorems I and II.

Proof of Theorem II. Let us assume that $f(x)$ is not zero almost everywhere. Since

$$g(u) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} f(x) e^{-iux}\, dx,$$

it follows for $w = u + iv$ that

$$|g(w)| \leqq (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} |f(x)| e^{vx}\, dx$$

if the right side exists. But for $v \geqq 0$, (2.01) gives

$$|g(w)| = O\left(\int_0^{\infty} e^{-h(x)+vx}\, dx + \int_{-\infty}^{0} |f(x)|\, dx\right)$$

or for $v \geqq 0$,

$$|g(w)| = O\left(\int_0^{\infty} e^{-h(x)+vx}\, dx\right). \tag{2.06}$$

$H(x)$ is the inverse function of $h'(x)$. Let $\xi = x - H(v)$. Then

$$h(x) - vx = h[\xi + H(v)] - v\xi - vH(v). \tag{2.07}$$

But by Taylor's theorem with remainder

$$h[\xi + H(v)] = h(H(v)) + \xi h'(H(v)) + \tfrac{1}{2}\xi^2 h''[H(v) + \alpha\xi],$$

where $0 < \alpha < 1$. Since $h'(H(v)) = v$, this becomes

$$h[\xi + H(v)] = h(H(v)) + \xi v + \tfrac{1}{2}\xi^2 h''[H(v) + \alpha\xi].$$

Using this in (2.07), we have

$$h(x) - vx = h(H(v)) - vH(v) + \tfrac{1}{2}\xi^2 h''[H(v) + \alpha\xi]. \tag{2.08}$$

By (2.02), $h''(x) \geqq c > 0$ and thus (2.08) becomes

$$-h(x) + vx \leqq -h(H(v)) + vH(v) - \tfrac{1}{2}c\xi^2.$$

Using this in (2.06), we get

$$g(w) = O\left(e^{-h(H(v))+vH(v)} \int_{-\infty}^{\infty} e^{-\frac{1}{2}c\xi^2}\, d\xi\right).$$

Or

$$g(w) = O(e^{vH(v)}) \qquad (v \geqq 0).$$

Since $H(v)$ is an increasing function, this in turn can be replaced by

$$g(w) = O(e^{vH(|w|)}) \qquad (v \geqq 0). \tag{2.09}$$

Thus $g(w)$ exists for $v \geqq 0$ and similarly $g'(w)$ exists for $v > 0$. This implies that $g(w)$ is analytic for $v > 0$. It is also clear that $g(w)$ is continuous for $v \geqq 0$. Since we have assumed that $f(x)$ is not equivalent to zero, it follows that $g(w)$ cannot be identically zero. Thus by a theorem of Carleman[4] there exists a constant A such that

$$\frac{1}{2\pi}\left(\int_{-R}^{-1} + \int_1^R\right)\left(\frac{1}{u^2} - \frac{1}{R^2}\right) \log |g(u)|\, du + \frac{1}{\pi R}\int_0^\pi \log |g(Re^{i\theta})| \sin\theta\, d\theta + A \geqq 0.$$

If we use (2.09), it follows that

$$\frac{1}{2\pi}\left(\int_{-R}^{-1} + \int_1^R\right)\left(\frac{1}{u^2} - \frac{1}{R^2}\right) \log |g(u)|\, du + \frac{1}{\pi R}\int_0^\pi R \sin^2\theta H(R)\, d\theta + A \geqq 0,$$

or

$$\frac{1}{2\pi}\left(\int_{-R}^{-1} + \int_1^R\right)\left(\frac{1}{u^2} - \frac{1}{R^2}\right) \log |g(u)|\, du + \tfrac{1}{2}H(R) + A \geqq 0.$$

But this contradicts (2.05). Thus $g(w) \equiv 0$, and this implies that $f(x)$ is equivalent to zero.

[4] E. C. Titchmarsh, *Theory of Functions*, Oxford, 1932, p. 130.

Proof of Theorem I. It is obvious that (2.05) is satisfied by (2.03) and (2.04) and thus Theorem I is an immediate consequence of Theorem II.

As regards (2.04), if we take the limiting case $h(x) = \infty$ in which case $H(x) = 0$, Theorem I reduces to the theorem containing (1.1), and (2.04) becomes (1.1) which is a best possible result.[5]

This justifies the integrated part of (2.04) but not the term $\pi H(x)$. That this term is necessary follows from

THEOREM III. *The constant* π *in* (2.04) *cannot be replaced by any smaller number.*

Proof of Theorem III. This theorem will be proved by giving an example. Let

$$f(x) = e^{-(1+ix)e^x}e^x.$$

Then clearly

$$h(x) = e^x - x$$

and thus $h'(x) = e^x - 1$. Therefore

$$H(x) = \log(1 + x).$$

Clearly

$$g(u) = (2\pi)^{-\frac{1}{2}}\int_{-\infty}^{\infty} e^{-(1+ix)e^x}e^{x-iux}\,dx.$$

Making a change of variable replacing e^x above by x, we have

$$g(u) = (2\pi)^{-\frac{1}{2}}\int_0^{\infty} e^{-x}x^{-ix-iu}\,dx$$

$$= (2\pi)^{-\frac{1}{2}}\int_0^{\infty} e^{-z}z^{-iz-iu}\,dz.$$

Rotating the path through an angle of $-\pi + \epsilon$, we have, setting $z = re^{i\theta}$,

$$|g(u)| \leqq (2\pi)^{-\frac{1}{2}}\int_0^{\infty} e^r e^{-r\log r\sin\epsilon + r(\pi-\epsilon)\cos\epsilon}e^{-u(\pi-\epsilon)}\,dr$$

$$= O(e^{-u(\pi-\epsilon)}) \qquad (u \to +\infty).$$

Thus $t(u) = (\pi - \epsilon)u$. Since ϵ is arbitrary, it is clear that if (2.04) involved $\pi - \delta$ instead of π then by taking $\epsilon = \frac{1}{2}\delta$ (2.04) would be satisfied and yet $f(x)$ is not equivalent to zero. This proves Theorem III.

3. Here we shall consider the second type of result relating the behavior of a function and its transform; that is, case (b), where $f(x)$ has a heavy zero for some finite $x = a$ and $g(u)$ tends to zero sufficiently fast at $+\infty$ (or at $-\infty$).

[5] See, for example, N. Wiener and R. E. A. C. Paley, *Fourier Transforms in the Complex Domain*, American Math. Soc. Coll. Pub., vol. XIX, Theorem XII.

It will be necessary for $f(x)$ to have this heavy zero only as a one-sided zero that is either for $a + 0$ or for $a - 0$. The theorem will be stated for the case $a + 0$, but $a - 0$ is of course equally good.

THEOREM IV. *For some a let*

$$f(x + a) = O(\exp\{-h(\log x^{-1})\}) \qquad (x \to +0), \tag{3.01}$$

where $h(x)$ satisfies the requirements of Theorem I. *Let $g(u) \in L(-\infty, \infty)$ and let $f(x)$ be its Fourier transform. Let*

$$e^{-\pi u} \int_0^u e^{\pi v} |g(v)| \, dv = O(e^{-t(u)}) \qquad (u \to \infty). \tag{3.02}$$

Then if $t(u)$ is an increasing function, if

$$t(u) < \pi u \qquad (u > 0), \tag{3.03}$$

and if $H(x)$ is the inverse function of $h'(x)$,

$$\int_1^x \frac{t(u)}{u^2} \, du - \pi H(x) \to \infty \qquad (x \to \infty) \tag{3.04}$$

implies that $f(x) \equiv 0$.

This theorem is slightly reminiscent of quasi-analyticity. However, since $f(x)$ need not possess even a first derivative, it is clear that this theorem covers a somewhat different situation from that of quasi-analyticity.

As an example of Theorem IV, let

$$f(x) = O(\exp\{-\exp[(\log x^{-1})^2]\}) \qquad (x \to +0).$$

Then $h(x) = e^{x^2}$ and therefore

$$H(x) = (\log x)^{\frac{1}{2}} + O(1) \qquad (x \to \infty).$$

Thus it suffices that

$$t(u) = \pi u (\log u)^{-\frac{1}{2}} + \theta(u),$$

where $\theta(u)$ satisfies (1.1), in order that (3.04) be satisfied and thus $f(x) \equiv 0$.

Proof of Theorem IV. With no loss of generality we can assume that $a = 0$. Let

$$F(x) = f(\log(1 + e^{-x})) \frac{e^{\frac{1}{2}x}}{(1 + e^x)^2}.$$

Since $g(u) \in L(-\infty, \infty)$, $f(x)$ is bounded and therefore $F(x) \in L(-\infty, \infty)$. Since, as $x \to +\infty$, $\log(1 + e^{-x}) \to +0$, it follows from (3.01) that

$$F(x) = O\left(\exp\left\{-h\left[\log \frac{1}{\log(1 + e^{-x})}\right]\right\}\right) \qquad (x \to +\infty). \tag{3.05}$$

But as $x \to \infty$

$$\log(1 + e^{-x}) < e^{-x}.$$

Thus since $h(x)$ is increasing,

$$h\left[\log \frac{1}{\log(1 + e^{-x})}\right] > h(\log e^{x}) = h(x).$$

Thus (3.05) gives

$$F(x) = O(e^{-h(x)}) \qquad (x \to +\infty). \tag{3.06}$$

Let $G(u)$ be the Fourier transform of $F(x)$. Then

$$\begin{aligned} G(u) &= (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} F(x) e^{-iux}\, dx \\ &= (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} f(\log(1 + e^{-x})) \frac{e^{-iux+\frac{1}{2}x}}{(1 + e^{x})^2}\, dx. \end{aligned}$$

Let $y = \log(1 + e^{-x})$. Then

$$x = \log\left(\frac{1}{e^{y} - 1}\right)$$

and thus

$$G(u) = (2\pi)^{-\frac{1}{2}} \int_0^{\infty} f(y)(e^{y} - 1)^{-iu+\frac{1}{2}} e^{-y}\, dy.$$

Since f is the Fourier transform of g,

$$\begin{aligned} G(u) &= \frac{1}{2\pi} \int_0^{\infty} (e^{y} - 1)^{-iu+\frac{1}{2}} e^{-y}\, dy \int_{-\infty}^{\infty} g(v) e^{iyv}\, dv \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} g(v)\, dv \int_0^{\infty} (e^{y} - 1)^{-iu+\frac{1}{2}} e^{-y+iyv}\, dy. \end{aligned}$$

Letting $e^{-y} = t$, we have

$$\begin{aligned} \int_0^{\infty} (e^{y} - 1)^{-iu+\frac{1}{2}} e^{-y+iyv}\, dy &= \int_0^{1} (1 - t)^{-iu+\frac{1}{2}} t^{-i(v-u)-\frac{1}{2}}\, dt \\ &= \frac{\Gamma(-iu + \frac{3}{2})\Gamma[-i(v - u) + \frac{1}{2}]}{\Gamma(-iv + 2)}. \end{aligned}$$

Thus

$$G(u) = \frac{1}{2\pi} \int_{-\infty}^{\infty} g(v) \frac{\Gamma(-iu + \frac{3}{2})\Gamma[-i(v - u) + \frac{1}{2}]}{\Gamma(-iv + 2)}\, dv.$$

If we use the Stirling formula for complex arguments of the Gamma function, it follows easily that for large u,

$$\begin{aligned} G(u) &= O\left(u \int_{-\infty}^{u} |g(v)| e^{-\frac{1}{2}\pi u - \frac{1}{2}\pi(u-v) + \frac{1}{2}\pi|v|}\, dv + u \int_u^{\infty} \frac{|g(v)|}{v^{\frac{3}{2}}}\, dv\right) \\ &= O\left(ue^{-\pi u} + ue^{-\pi u} \int_0^{u} |g(v)| e^{\pi v}\, dv + u \int_u^{\infty} \frac{|g(v)|}{v^{\frac{3}{2}}}\, dv\right) \qquad (u \to +\infty). \end{aligned}$$

By (3.02) this becomes

$$(3.07)\qquad G(u) = O\left(ue^{-\pi u} + ue^{-t(u)} + u\int_u^\infty \frac{g(v)}{v^3}\,dv\right) \qquad (u \to +\infty).$$

Clearly (3.02) implies

$$e^{-\pi u}\int_{u-1}^u e^{\pi v}\,|\,g(v)\,|\,dv = O(e^{-t(u)}),$$

and therefore

$$\int_{u-1}^u |\,g(v)\,|\,dv = O(e^{-t(u)}) \qquad (u \to +\infty).$$

This in turn implies that for all $k \geqq 0$,

$$(3.08)\qquad \int_{u+k}^{u+k+1} |\,g(v)\,|\,dv = O(e^{-t(u+k+1)}) = O(e^{-t(u)}) \qquad (u \to +\infty).$$

Thus

$$\int_u^\infty \frac{g(v)}{v^3}\,dv = O(e^{-t(u)})\sum_{k=0}^\infty \frac{1}{(u+k)^3} = O\left(\frac{1}{u^2}e^{-t(u)}\right).$$

Using this in (3.07), we have

$$G(u) = O(ue^{-\pi u} + ue^{-t(u)}) \qquad (u \to +\infty).$$

By (3.03) $t(u) < \pi u$ and thus

$$(3.09)\qquad G(u) = O(ue^{-t(u)}) \qquad (u \to +\infty).$$

We now apply Theorem II to $F(x)$. Since $F(x)$ satisfies (3.06), and $G(u)$ satisfies (3.09), and since $G(u)$ is bounded for $u < 0$, (2.05) of Theorem II becomes

$$\varliminf_{x\to\infty}\left\{\int_1^x\left(\frac{1}{y^2} - \frac{1}{x^2}\right)[\log y - t(y)]\,dy + \pi H(x)\right\} = -\infty,$$

or

$$\varliminf_{x\to\infty}\left\{-\int_1^x\left(\frac{1}{y^2} - \frac{1}{x^2}\right)t(y)\,dy + \pi H(x)\right\} = -\infty.$$

Since $t(y) < \pi y$, it is clear that the above result is implied by (3.04). Thus by Theorem II, $F(x)$ is zero almost everywhere. But from the definition of $F(x)$ this implies that $f(x) \equiv 0$ $(x \geqq 0)$. But this means that

$$(3.10)\qquad f(x+1) = O\left(\exp\left\{-h\left(\log\frac{1}{|\,x\,|}\right)\right\}\right) \qquad (x \to -0).$$

If we apply the result of the proof of Theorem IV obtained up to this point to (3.10), it is clear that we obtain $f(x) \equiv 0$ $(x \leqq 1)$. But this with the result already obtained completes the proof of Theorem IV.

4. Here we shall improve the constant π in (2.04) for the case where $h(x)$ is of the form ax^p. By elaborating the formal aspects of the proof, the case $h(x) = ax^p(\log x)^{m_1}(\log\log x)^{m_2}$ etc. can also be covered.

THEOREM V. *Let $f(x) \in L(-\infty, \infty)$ and let*

$$f(x) = O(e^{-ax^p}) \qquad (x \to +\infty), \tag{4.01}$$

where $p \geqq 2$. Let

$$a' = \frac{1}{p'}\left(\frac{1}{ap}\right)^{p'/p} \sin\frac{\pi}{p}, \tag{4.02}$$

where

$$\frac{1}{p'} = 1 - \frac{1}{p}.$$

Let $g(u)$, the Fourier transform of $f(x)$, satisfy

$$g(u) = O(\exp[-a'u^{p'} - \theta(u^{p'})]) \qquad (u \to +\infty), \tag{4.03}$$

where $\theta(u)$ is increasing and satisfies (1.1). *Then $f(x)$ is zero almost everywhere.*

If $p = 2$ and $a = \frac{1}{2}$, then $a' = \frac{1}{2}$ and thus in this case Theorem V is satisfactory. However, if $p > 2$ then the constants are not as good as those in Morgan's theorem, but it seems quite plausible that this is the case because of the considerably weaker restriction imposed by the one-sided conditions used here.

Proof of Theorem V. If we set $h(x) = ax^p$, then (4.01) gives

$$f(x) = O(e^{-h(x)}) \qquad (x \to +\infty).$$

Proceeding now exactly as in the proof of Theorem II, we obtain the formula several lines below (2.08),

$$g(w) = O(e^{-h(H(v))+vH(v)}) \qquad (v \geqq 0).$$

Here $h'(x) = apx^{p-1}$ and thus

$$H(v) = \left(\frac{v}{ap}\right)^{1/(p-1)}.$$

Thus

$$\begin{aligned} g(w) &= O\left(\exp\left[-a\left(\frac{v}{ap}\right)^{p/(p-1)} + \left(\frac{1}{ap}\right)^{1/(p-1)} v^{1+1/(p-1)}\right]\right) \\ &= O\left(\exp\left[v^{p'}\left(\frac{1}{ap}\right)^{1/(p-1)}\left(1 - \frac{1}{p}\right)\right]\right). \end{aligned}$$

Or

$$g(w) = O\left(\exp\left[v^{p'}\left(\frac{1}{ap}\right)^{p'/p}\frac{1}{p'}\right]\right) \qquad (v \geqq 0). \tag{4.04}$$

Let $w = s^{1/p'}$. Let

$$g(w) = g(s^{1/p'}) = G(s).$$

Then for $s = \sigma + i\tau$

$$G(-\sigma) = O\left(\exp\left[\sigma \sin\frac{\pi}{p'}\left(\frac{1}{ap}\right)^{p'/p}\frac{1}{p'}\right]\right) \qquad (\sigma \to +\infty).$$

Or

$$G(-\sigma) = O(e^{a'\sigma}) \qquad (\sigma \to +\infty). \tag{4.05}$$

But by (4.03)

$$G(\sigma) = O(e^{-a'\sigma-\theta(\sigma)}) \qquad (\sigma \to +\infty). \tag{4.06}$$

Let

$$T(s) = e^{a's}G(s).$$

Then by (4.05) and (4.06), $T(s)$ is bounded on the real axis and

$$T(\sigma) = O(e^{-\theta(\sigma)}) \qquad (\sigma \to +\infty). \tag{4.07}$$

Also by (4.04)

$$T(s) = O(e^{a'|s|/\sin(\pi/p)}) \qquad (\tau \geqq 0). \tag{4.08}$$

If we now assume $T(s) \not\equiv 0$ and apply Carleman's theorem[6] to $T(s)$ in the upper half-plane, and if we use (4.08) and the fact that $T(\sigma)$ is bounded, we have

$$\int_1^R \log|T(\sigma)|\left(\frac{1}{\sigma^2} - \frac{1}{R^2}\right)d\sigma > -A,$$

where A is some constant. If we use (4.07), this becomes

$$\int_1^R \left(\frac{1}{\sigma^2} - \frac{1}{R^2}\right)\theta(\sigma)\,d\sigma < A.$$

Or

$$\int_1^{\frac{1}{2}R} \left(\frac{1}{\sigma^2} - \frac{1}{R^2}\right)\theta(\sigma)\,d\sigma < A.$$

But for $1 < \sigma < \frac{1}{2}R$, $R^{-2} < \frac{1}{4}\sigma^{-2}$. Thus

$$\int_1^{\frac{1}{2}R} \frac{\theta(\sigma)}{\sigma^2}\,d\sigma < 2A.$$

If we let $R \to \infty$, this contradicts (1.1). Thus $T(s) \equiv 0$. But this means $g(u) \equiv 0$ and therefore $f(x)$ is zero almost everywhere.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY.

[6] Titchmarsh, loc. cit.

NUMERICAL ANALYSIS OF CERTAIN FREE DISTRIBUTIVE STRUCTURES

By Randolph Church

Consider the set Σ_n of all formal cross-cuts and unions[1] of n symbols $A_1, A_2, \cdots, A_n$. Disjoint classes which exhaust Σ_n can be formed with respect to an equivalence introduced according to the axioms of a distributive structure if a suitable axiom as to the independence of the A_i is assumed. A decision as to the equality of the classes containing arbitrary elements of Σ_n can be reached in a finite number of steps. These classes form the elements of a distributive structure, Δ_n, the *free distributive structure based on n elements.* Its elements can be represented by the unique cross-cut of unions of the A_i contained in each class. Δ_n contains a finite number of elements, $N(\Delta_n)$, its order.[2]

Dedekind[3] gave the order of Δ_n for $n \leqq 4$. The purpose of this paper is to present an analysis of $N(\Delta_n)$, $n \leqq 5$. The analysis depends on the notion of conjugate elements. Let X_1 and X_2 be two elements of Δ_n, written as the cross-cut of unions of the A_i; if there exists a permutation α of the A_i such that $\alpha X_1 = X_2$, we say that X_2 is conjugate to X_1. The relation of conjugacy is symmetric, reflexive and transitive, dividing Δ_n into disjoint sets $\{X\}$ of conjugate elements. The number of conjugates in a set $\{X\}$ is $h = n!/k$, where k is the order of $G_n(X)$, the group of degree n which leaves X unchanged. A conjugate belongs to a transformed group: $G_n(\alpha X) = \alpha G_n(X)\alpha^{-1}$. The rank of an element in Δ_n is invariant under permutations of the A_i so that the elements of a set of conjugates are of the same rank.

The facts thus sketched determine the arrangement of the following tables. The number of elements of rank r, denoted by N_r, is given in the right-hand column, so that the sum of the entries in this column is $N(\Delta_n)$. The entries in the body of a table give, for each value of r, the number of sets consisting of h conjugates. The data presented here was obtained by listing representatives

Received April 29, 1940; presented to the American Mathematical Society, April 26, 1940.

[1] For the axioms and essential properties of distributive structures, reference may be made to O. Ore, *On the foundations of abstract algebra*, I, Annals of Math., vol. 36(1935), pp. 406–437.

[2] The details of this existence proof were included in the writer's dissertation, Yale, 1935. The chains of Boolean structures (defective with respect to one unit) composing the free distributive structure, referred to at the end of this paper, were there considered in detail.

[3] R. Dedekind, *Über Zerlegungen von Zahlen durch ihre grössten gemeinsamen Teiler*, Werke II, Braunschweig, 1931, pp. 103–147; p. 147.

$n = 2$

$r \setminus h$	1	2	N_r
2	1		1
1		1	2
0	1		1
	2	1	4

$n = 3$

$r \setminus h$	1	3	N_r
6	1		1
5		1	3
4		1	3
3	1	1	4
2		1	3
1		1	3
0	1		1
	3	5	18

$n = 4$

$r \setminus h$	1	3	4	6	12	N_r
14	1					1
13			1			4
12				1		6
11			1	1		10
10	1				1	13
9				1	1	18
8		1	1		1	19
7			3		1	24
6		1	1		1	19
5				1	1	18
4	1				1	13
3			1	1		10
2				1		6
1			1			4
0	1					1
	4	2	9	6	7	166

$n = 5$

$r \setminus h$	1	5	10	12	15	20	30	60	120	N_r
30	1									1
29		1								5
28			1							10
27			2							20
26		1					1			35
25	1						2			61
24			2		1			1		95
23			1		1	2	1	1		155
22			1		1	2	1	2		215
21		1	2		1		1	4		310
20		1	1	1			2	5		387
19		1	1		1	1	2	6		470
18			1			2	4	4	1	530
17					2	2	3	5	1	580
16		1	1			1	3	6	1	605
15	1	2	2			1	3	6	1	621
14		1	1			1	3	6	1	605
13					2	2	3	5	1	580
12			1			2	4	4	1	530
11		1	1		1	1	2	6		470
10		1	1	1			2	5		387
9		1	2		1		1	4		310
8			1		1	2	1	2		215
7			1		1	2	1	1		155
6			2		1			1		95
5	1						2			61
4		1					1			35
3			2							20
2			1							10
1		1								5
0	1									1
	5	14	28	2	14	21	43	74	7	7579

of the classes of conjugates together with the number of conjugates in each class and determining the corresponding ranks by obtaining the structure inclusion relations between the elements of the various classes. The sets of conjugates thus obtained were checked by converting the representative elements into their duals, and by arranging the elements of the structure in chains of Boolean structures according to general theory.

It is considered desirable to make the results here given available now instead of including them with related investigations as originally planned when the analysis of Δ_5 was completed in 1936.

UNITED STATES NAVAL ACADEMY.

LINEAR FORMS AND POLYNOMIALS IN A GALOIS FIELD

By L. Carlitz

1. **Introduction.** Let $t, u_0, u_1, \cdots$ be indeterminates over the Galois field $GF(p^n)$, and put

$$(1.1) \qquad f_m(t) = f_m(t; u_0, \cdots, u_{m-1}) = \prod_{(c)} (t + c_0 u_0 + \cdots + c_{m-1} u_{m-1}),$$

the product extending over all sets $(c_0, \cdots, c_{m-1})$ in $GF(p^n)$. Then by a formula due to E. H. Moore[1] we may express $f_m(t)$ as the quotient of two determinants:

$$(1.2) \qquad f_m(t; u_0, \cdots, u_{m-1}) = \frac{D(u_0, \cdots, u_{m-1}, t)}{D(u_0, \cdots, u_{m-1})},$$

where

$$(1.3) \qquad D_m = D(u_0, \cdots, u_m) = \begin{vmatrix} u_0 & u_0^{p^n} & \cdots & u_0^{p^{nm}} \\ \cdots & \cdots & \cdots & \cdots \\ u_m & u_m^{p^n} & \cdots & u_m^{p^{nm}} \end{vmatrix},$$

and $D(u_0, \cdots, u_{m-1}, t)$ is defined by taking $u_m = t$.

In the special case[2]

$$(1.4) \qquad u_i = x^i \qquad (i = 0, 1, 2, \cdots),$$

where x is an indeterminate, the linear form

$$c_0 u_0 + c_1 u_1 + \cdots + c_{m-1} u_{m-1} \qquad (c_j \text{ in } GF(p^n))$$

reduces to $c_0 + c_1 x + \cdots + c_{m-1} x^{m-1}$, the general polynomial of degree $< m$, and $f_m(t)$ reduces to $\psi_m(t)$. We shall refer to (1.4) as the P-case. The chief object of the present paper is to extend certain properties of $\psi_m(t)$ to the more general $f_m(t)$. In particular we define an operator Ω^i for which

$$g(tw) = \sum \Omega^i g(w) f_i(t),$$

where $g(t)$ is an arbitrary linear[3] polynomial in t. Applications are made to the evaluation of certain power sums. The inverse of $f_m(t)$ and certain limiting cases are discussed briefly. Finally we define a polynomial $G_k(t)$ of degree k that reduces to $f_m(t)$ for $k = p^{nm}$.

Received May 8, 1940.

[1] Bulletin of the American Mathematical Society, vol. 2(1896), pp. 189–195. See also O. Ore, Transactions of the American Mathematical Society, vol. 35(1933), pp. 559–584; L. E. Dickson, Trans. Amer. Math. Soc., vol. 12(1911), p. 75.

[2] This Journal, vol. 1(1935), pp. 137–168. Cited as I.

[3] That is, of the form $\sum \beta_i t^{p^{ni}}$.

2. **Preliminary formulas.** Generalizing (1.3) we define[4]

$$[i_0, i_1, \cdots, i_m] = \begin{vmatrix} u_0^{p^{ni_0}} & u_0^{p^{ni_1}} & \cdots & u_0^{p^{ni_m}} \\ \cdots & \cdots & \cdots & \cdots \\ u_m^{p^{ni_0}} & u_m^{p^{ni_1}} & \cdots & u_m^{p^{ni_m}} \end{vmatrix}, \tag{2.1}$$

so that in particular

$$[0, 1, \cdots, m] = D(u_0 u_1 \cdots u_m) = D_m .$$

Then (1.2) becomes

$$f_m(t) = \sum_{i=0}^{m} (-1)^{m-i} A_i^m t^{p^{ni}}, \tag{2.2}$$

where

$$A_i^m = \frac{[0, \cdots, i-1, i+1, \cdots, m]}{D_{m-1}}. \tag{2.3}$$

It is clear from the definition that A_i^m is integral, that is, a polynomial in $u_0, \cdots, u_{m-1}$; this may also be proved directly[5] from (2.3). Note in particular that

$$A_0^m = D_{m-1}^{p^n-1}, \qquad A_m^m = 1.$$

It will be convenient to define another polynomial

$$\varphi_m(t) = (-1)^m \frac{D(u_0 \cdots u_{m-1} t)}{D_{m-1}^{p^n}} = \frac{(-1)^m f_m(t)}{D_{m-1}^{p^n-1}}; \tag{2.4}$$

so that the coefficient of t in $\varphi_m(t)$ is 1. Evidently (1.1) implies

$$\varphi_m(t) = t \prod{}' \left(1 + \frac{t}{c_0 u_0 + \cdots + c_{m-1} u_{m-1}}\right), \tag{2.5}$$

where in the product the set $(0, \cdots, 0)$ is omitted.

Consider now $f_{m+1}(t) = f_{m+1}(t; u_0, \cdots, u_m)$. By (1.1) we have

$$\begin{aligned} f_{m+1}(t) = \prod_{c_m} \prod_{c_0, \cdots, c_{m-1}} & (t + c_0 u_0 + \cdots + c_{m-1} u_{m-1} + c_m u_m) \\ &= \prod_c f_m(t + c u_m) = \prod_c \{f_m(t) + c f_m(u_m)\}, \end{aligned}$$

and therefore

$$f_{m+1}(t) = f_m^{p^n}(t) - f_m^{p^n-1}(u_m) f_m(t), \tag{2.6}$$

where by (1.2)

$$f_m(u_m) = \frac{D(u_0 \cdots u_{m-1} u_m)}{D(u_0 \cdots u_m)} = \frac{D_m}{D_{m-1}}.$$

[4] See D. E. Rutherford, *Modular Invariants*, Cambridge, 1932, p. 15.

[5] Rutherford, loc. cit.

Clearly $f_m(u_m)$ is integral. Note that in the P-case $f_m(u_m)$ becomes

$$F_m = (x^{p^{nm}} - x)(x^{p^{nm}} - x^{p^n}) \cdots (x^{p^{nm}} - x^{p^{n(m-1)}}).$$

It follows immediately from (1.2) that, for m fixed, $f_m(x)$ remains unchanged under the transformation

$$u_i = \sum_{j=0}^{i} c_{ij} v_j, \qquad | c_{ij} | \neq 0.$$

On the other hand the set $f_0(t), f_1(t), \cdots, f_m(t)$ will remain unchanged only under the more restricted transformation

$$u_i = \sum_{j=0}^{i} c_{ij} v_j, \qquad \prod_{j=0}^{m-1} c_{jj} \neq 0, \tag{2.7}$$

as is easily proved by induction.

Note also that

$$f_m(tv; u_0 v, \cdots, u_{m-1} v) = v^{p^{nm}} f_m(t; u_0, \cdots, u_{m-1}),$$

as follows from (1.1), while by (2.5) the corresponding formula for $\varphi_m(t)$ is simply

$$\varphi_m(tv; u_0 v, \cdots, u_{m-1} v) = \varphi_m(t; u_0, \cdots, u_{m-1}),$$

so that $f_m(t)$ is homogeneous of degree p^{nm} in the m variables $t, u_0, \cdots, u_{m-1}$, and $\varphi_m(t)$ is homogeneous of degree 0. Thus there is little loss in generality in taking $u_0 = 1$.

3. **An expansion formula.** Let

$$g(t) = \sum_{i=0}^{k} \alpha_i t^{p^{ni}}$$

be any linear polynomial in t; we shall call k the *order* of $g(t)$. Clearly we may put

$$g(t) = \sum_{i=0}^{k} C_i f_i(t);$$

more generally we may put

$$g(tw) = \sum_{i=0}^{k} C_i(w) f_i(t), \tag{3.1}$$

where $C_i(w)$ is a linear polynomial in the indeterminate w. In order to determine $C_i(w)$ we define an operator Ω^i as follows:

$$\Omega g(t) = \Omega^1 g(t) = \frac{u_0 g(u_1 t) - u_1 g(u_0 t)}{D(u_0 u_1)},$$

$$\Omega^2 g(t) = \frac{D(u_0 u_1) g(u_2 t) - D(u_0 u_2) g(u_1 t) + D(u_1 u_2) g(u_0 t)}{D(u_0 u_1 u_2)},$$

and generally

$$(3.2) \qquad \Omega^i g(t) = \frac{D(u_0 \cdots u_{i-1})g(u_i t) - \cdots + (-1)^i D(u_1 \cdots u_i)g(u_0 t)}{D(u_0 u_1 \cdots u_i)}.$$

For completeness we put

$$\Omega^0 g(t) = \frac{g(u_0 t)}{u_0},$$

which is simply $g(t)$ for $u_0 = 1$. In particular we may verify that for $g(t) = t^{p^{nk}}$ we get

$$(3.3) \qquad \Omega^i t^{p^{nk}} = \frac{[0, \cdots, i-1, k]}{D_i} t^{p^{nk}}.$$

Note that (3.3) implies

$$\Omega^i t^{p^{nk}} = \begin{cases} 0 & \text{for } i > k, \\ t^{p^{nk}} & \text{for } i = k; \end{cases}$$

in other words, Ω^i annihilates linear polynomials of order $< i$. Note also that for $g(t) = f_k(t)$,

$$(3.4) \qquad \Omega^i f_k(1) = \begin{cases} 0 & \text{for } i < k, \\ 1 & \text{for } i = k, \end{cases}$$

by direct substitution in (3.2).

Returning to (3.1), put $t = u_0$ and we get

$$C_0(w) = \Omega^0 g(w).$$

Next apply Ω to both members of (3.1), thought of as functions of t:

$$\Omega g(tw) = \sum_{i=1}^{k} C_i(w)\Omega f_i(t);$$

then (3.4) implies

$$C_1(w) = \Omega g(w).$$

Generally, apply Ω^i to (3.1):

$$\Omega^i g(tw) = \sum_{j=i}^{k} C_j(w)\Omega^i f_j(t),$$

and (3.4) implies

$$C_i(w) = \Omega^i g(w).$$

Thus (3.1) becomes

$$(3.5) \qquad g(tw) = \sum_{i=0}^{k} \Omega^i g(w) f_i(t).$$

In the P-case Ω^i reduces to Δ^i/F_i, where Δ^i may be defined recursively by[6]

$$\Delta^{i+1}g(t) = \Delta^i g(xt) - x^{p^{ni}}\Delta^i g(t).$$

In this case there is the identity

$$\Delta\psi_m(t) = (x^{p^{nm}} - x)\psi_{m-1}^{p^n}(t),$$

which is indeed characteristic of the P-case. For take $u_0 = 1$, and assume

$$\Omega f_m(t) = B_m f_{m-1}^{p^n}(t), \tag{3.6}$$

where B_m is independent of t, and m takes on all values (or all values not exceeding some k). Since by (1.1)

$$f_{m-1}(u_i) = 0 \qquad (i = 0, \cdots, m-2),$$

it follows from (3.6) that

$$f_m(u_1u_i) = f_m(u_i) = 0,$$

and therefore in particular

$$u_1u_{m-2} = \sum_{j=0}^{m-1} c_{m-1,j}u_j \qquad (m = 2, 3, \cdots).$$

From this it follows that either

(i) u_1 is algebraic over $GF(p^n)$, or

(ii) $u_m = \sum_{j=0}^{m} b_{mj}u_1^j$.

Case (i) will occur if some $c_{ii} = 0$; case (ii) will occur if no $c_{ii} = 0$. Clearly (ii) implies the P-case.

As a particular instance of (3.5) take $g(t) = t^{p^{nk}}$. Then applying (3.3) we get the identity

$$t^{p^{nk}} = \sum_{i=0}^{k} \frac{[0, \cdots, i-1, k]}{D_i} f_i(t), \tag{3.7}$$

which will be used later.

4. **An identity.** Let us consider the following[7]

PROBLEM. To determine C_i such that

$$\varphi_m + C_1\varphi_m^{p^n} + \cdots + C_{k-1}\varphi_m^{p^{n(k-1)}} = t - \sum_{j=0}^{m-1} B_{kj}t^{p^{n(k+j)}}, \tag{4.1}$$

where C_i and B_{kj} are independent of t, and $\varphi_m(t)$ is defined by (2.4). That an identity of the form (4.1) exists is fairly clear. Indeed by adding $B_{k0}\varphi_m^{p^{nk}}$ to both members of (4.1) we eliminate the term in $t^{p^{nk}}$ on the right, so that we are passing from the case k to the case $k+1$. Incidentally this proves

$$C_k = B_{k0}. \tag{4.2}$$

[6] See I, p. 143.

[7] Compare this Journal, vol. 5(1939), pp. 941–947, p. 944. Cited as II.

As for B_{kj}, put $t = u_i$ in (4.1):

$$\sum_{j=0}^{m-1} B_{kj} u_i^{p^{n(k+j)}} = u_i \qquad (i = 0, \cdots, m-1). \tag{4.3}$$

Fix k, then (4.3) is a system of m equations in the m quantities B_{kj}. The denominator is evidently

$$| u_i^{p^{n(k+j)}} | = D^{p^{nk}}(u_0 \cdots u_{m-1}) = D_{m-1}^{p^{nk}}$$

and thus

$$\begin{aligned} D_{m-1}^{p^{nk}} B_{kj} &= [k, \cdots, k+j-1, 0, k+j+1, \cdots, k+m-1] \\ &= (-1)^j [0, k, \cdots, k+j-1, k+j+1, \cdots, k+m-1]. \end{aligned} \tag{4.4}$$

For $j = 0$ this becomes

$$D_{m-1}^{p^{nk}} B_{k0} = [0, k+1, \cdots, k+m-1]. \tag{4.5}$$

This together with (4.2) explicitly determines C_i.

Various identities similar to (4.1) may be derived by the same method. In the right member of (4.1) the terms in

$$t^{p^n}, t^{p^{2n}}, \cdots, t^{p^{n(k-1)}}$$

are missing; we may eliminate other groups of $k - 1$ terms. We may still compute explicitly the corresponding B_{kj}, but it is more difficult to determine C_i.

Let us take in particular

$$\varphi_m + \gamma_1 \varphi_m^{p^n} + \cdots + \gamma_m \varphi_m^{p^{nm}} = \sum_{j=0}^{m} \beta_j t^{p^{2nj}}, \tag{4.6}$$

where $\beta_0 = 1$. Put $t = u_i$ and (4.6) becomes

$$\sum_{j=1}^{m} \beta_j u_i^{p^{2nj}} = -u_i \qquad (i = 0, \cdots, m-1). \tag{4.7}$$

The denominator of β_j

$$= | u_i^{p^{2nj}} | = \bar{D}^{p^{2n}}(u_0 \cdots u_{m-1}) = \bar{D}_{m-1}^{p^{2n}},$$

the bar on D indicating that in the definition (1.3) we are replacing n by $2n$; in other words, the underlying Galois field is $GF(p^{2n})$. Thus we get

$$\begin{aligned} \bar{D}_{m-1}^{p^{2n}} \beta_j &= -[2, \cdots, 2j-2, 0, 2j+2, \cdots, 2m] \\ &= (-1)^j [0, 2, \cdots, 2j-2, 2j+2, \cdots, 2m], \end{aligned}$$

and therefore by (2.3),

$$\bar{D}_{m-1}^{p^{2n}-1} \beta_j = (-1)^j \bar{A}_j^m.$$

Comparison with (2.4) now gives

$$\sum_{j=0}^{m} \beta_j t^{p^{2nj}} = \bar{\varphi}_m(t),$$

where $\bar{\varphi}_m$ is defined relative to $GF(p^{2n})$. Thus we have explicitly determined the right member of (4.6). The result for γ_i is not simple and will be omitted.

As another identity similar to (4.1) we cite

$$\sum_{j=0}^{k} A_{kj} f_m^{p^{nj}} = t^{p^{n(k+m)}} - \sum_{j=0}^{m-1} B_{kj} t^{p^{nj}}, \tag{4.8}$$

where $A_{kk} = 1$. Raising both members to the p^n-th power, we have at once

$$A_{k+1,0} = B_{k,m-1}^{p^n}, \qquad A_{k+1,j+1} = A_{kj}^{p^n},$$

and from this it follows that

$$A_{k+1,j} = A_{k+1-j,0}^{p^{nj}} = B_{k-j,m-1}^{p^{n(j+1)}}.$$

It remains to determine B_{kj}. In (4.8) put $t = u_i$; this gives

$$\sum_{j=0}^{m-1} B_{kj} u_i^{p^{nj}} = u_i^{p^{n(k+m)}} \qquad (i = 0, \cdots, m-1). \tag{4.9}$$

Solving (4.9) we get

$$D_{m-1}B_{kj} = [0, \cdots, j-1, k+m, j+1, \cdots, m-1],$$

and in particular

$$D_{m-1}B_{k,m-1} = [0, \cdots, m-2, k+m]. \tag{4.10}$$

This result is useful in certain applications.

5. **Another expansion formula.** Analogous to the expansion of §3 we may consider

$$g(t) = \sum_{i=0}^{m} C_i \varphi_i^{p^{n(m-i)}}(t), \tag{5.1}$$

where

$$g(t) = \sum_{i=0}^{k} \alpha_i t^{p^{ni}} \qquad (m \geqq k).$$

It is fairly clear that the representation (5.1) is possible. To determine C_i we consider the more general

$$g(tw) = \sum_{i=0}^{m} C_i(w) \varphi_i^{p^{n(m-i)}}(t), \tag{5.2}$$

where $C_i(w)$ is a linear polynomial.

Define an operator $O^{m,j}$ by means of

$$O^{m,0}g(t) = \frac{g(u_0 t)}{u_0^{p^{nm}}},$$

so that for $u_0 = 1$, $O^{m,0}g(t) = g(t)$,

$$O^{m,1}g(t) = -\frac{u_0^{p^{nm}}g(u_1 t) - u_1^{p^{nm}}g(u_0 t)}{D_1^{p^{n(m-1)}}},$$

$$O^{m,2}g(t) = \frac{D^{p^{n(m-1)}}(u_0 u_1)g(u_2 t) - D^{p^{n(m-1)}}(u_0 u_2)g(u_1 t) + D^{p^{n(m-1)}}(u_1 u_2)g(u_0 t)}{D_2^{p^{n(m-2)}}},$$

and generally

$$O^{m,j}g(t) = (-1)^j$$

$$\cdot \frac{D^{p^{n(m-j+1)}}(u_0 \cdots u_{j-1})g(u_j t) - \cdots + (-1)^j D^{p^{n(m-j+1)}}(u_1 \cdots u_j)g(u_0 t)}{D_j^{p^{n(m-j)}}}. \tag{5.3}$$

If we take $g(t) = t^{p^{nk}}$, we get

$$O^{m,j}t^{p^{nk}} = (-1)^j \frac{[m-j+1, \cdots, m, k]}{D_j^{p^{n(m-j)}}} t^{p^{nk}}, \tag{5.4}$$

so that in particular

$$O^{m,j}t^{p^{nk}} = \begin{cases} 0 & \text{for } m-j+1 \leqq k \leqq m, \\ t^{p^{nk}} & \text{for } m-j = k. \end{cases} \tag{5.5}$$

As a consequence of (5.5) we have

$$O^{m,j}\varphi_i^{p^{n(m-i)}}(t) = 0 \qquad \text{for } i < j, \tag{5.6}$$

and by (5.3),

$$O^{m,j}\varphi_i^{p^{n(m-i)}}(1) = \begin{cases} 0 & \text{for } i > j, \\ 1 & \text{for } i = j. \end{cases} \tag{5.7}$$

Turn now to (5.2); application of $O^{m,j}$ gives

$$O^{m,j}g(tw) = \sum_{i=j}^{m} C_i(w)O^{m,j}\varphi_i^{p^{n(m-i)}}(t),$$

since by (5.6) the terms for which $i < j$ vanish. Now put $t = 1$ and use (5.7); this gives at once

$$C_i(w) = O^{m,i}g(w). \tag{5.8}$$

Substituting from (5.8) into (5.2), we get finally

$$g(tw) = \sum_{i=0}^{m} O^{m,i}g(w)\cdot\varphi_i^{p^{n(m-i)}}(t). \tag{5.9}$$

As an application of (5.9), take $g(t) = t^{p^{nk}}$; then by (5.4) and (5.5) we have

$$t^{p^{nk}} = \sum_{i=0}^{m} (-1)^i \frac{[m-i+1, \cdots, m, k]}{D_i^{p^{n(m-i)}}} \varphi_i^{p^{n(m-i)}}(t)$$

$$= \sum_{i=0}^{m-k} \frac{[k, m-i+1, \cdots, m]}{D_i^{p^{n(m-i)}}} \varphi_i^{p^{n(m-i)}}(t).$$

Replacing $m - k$ by m, we see that this reduces to the relation

$$t = \sum_{i=0}^{m} \frac{[0, m-i+1, \cdots, m]}{D_i^{p^{n(m-i)}}} \varphi_i^{p^{n(m-i)}}(t), \tag{5.10}$$

which may be compared with (3.7).

In the P-case (5.10) reduces to the (new) formula

$$t = \sum_{j=0}^{m} \frac{L_m}{L_{m-j} L_j^{p^{n(m-j)}}} \pi_j^{p^{n(m-j)}}(t),$$

where

$$L_m = (x^{p^{nm}} - x)(x^{p^{n(m-1)}} - x) \cdots (x^{p^n} - x), \tag{5.11}$$

$$\pi_m(t) = (-1)^m \frac{L_m}{F_m} \psi_m(t).$$

We remark that L_m is the P-form of $D_m/D_{m-1}^{p^n}$ (compare (6.3) below) which is not integral in the general case; on the other hand, as we have seen above, F_m is the reduced form of D_m/D_{m-1}, which is integral.

6. **Power sums.** We require the following lemma,[8] which is easily proved by induction: Let

$$g(t) = \prod_{i=1}^{k} (t + \gamma_i) = t^k + \beta_1 t^{k-1} + \cdots + \beta_k,$$

$$g'(t) = kt^{k-1} + (k-1)\beta_1 t^{k-2} + \cdots + \beta_{k-1};$$

then

$$\frac{g'(t)}{g(t)} = \sum_{i=1}^{k} \frac{1}{t + \gamma_i}. \tag{6.1}$$

In this identity take $g(t) = f_m(t)$. Since by (1.2) $f'_m(t)$ reduces to $(-1)^m D^{p^n-1}$, we get

$$\frac{(-1)^m D_{m-1}^{p^n-1}}{f_m(t)} = \sum_{(c)} \frac{1}{t + c_0 u_0 + \cdots + c_{m-1} u_{m-1}}. \tag{6.2}$$

[8] I, p. 160.

This relation may be put in the form

$$\sum_{(c)} \frac{1}{c_0 u_0 + \cdots + c_{m-1} u_{m-1} + u_m} = (-1)^m \frac{D^{p^n}(u_0 \cdots u_{m-1})}{D(u_0 \cdots u_m)} \tag{6.3}$$
$$= (-1)^m \frac{D_{m-1}^{p^n}}{D_m}.$$

We shall use the abbreviation

$$U_m = c_0 u_0 + \cdots + c_{m-1} u_{m-1}, \qquad W_m = U_m + u_m.$$

Consider now the sum

$$\sum_{(c)} \frac{f_k(t + U_m)}{t + U_m} \tag{6.4}$$

for $k < m$. If we use (6.2), it is clear that (6.4)

$$= \sum_{c_k, \cdots, c_{m-1}} \sum_{c_0, \cdots, c_{k-1}} \frac{f_k(t + c_k u_k + \cdots + c_{m-1} u_{m-1})}{t + c_k u_k + \cdots + c_{m-1} u_{m-1} + U_k}$$
$$= \sum_{c_k, \cdots, c_{m-1}} (-1)^k D_{k-1}^{p^n - 1} \frac{f_k(t + c_k u_k + \cdots + c_{m-1} u_{m-1})}{f_k(t + c_k u_k + \cdots + c_{m-1} u_{m-1})} = 0.$$

For $k > m$, (6.4) is obviously 0, while for $k = m$, we get by (6.2)

$$\frac{(-1)^m D_{m-1}^{p^n - 1}}{f_m(t)} \cdot f_m(t) = (-1)^m D_{m-1}^{p^n - 1}.$$

Combining the several cases and replacing t by u_m, we have finally[9]

$$\sum_{(c)} \frac{f_k(W_m)}{W_m} = \begin{cases} 0 & \text{for } k \neq m, \\ (-1)^m D_{m-1}^{p^n - 1} & \text{for } k = m. \end{cases} \tag{6.5}$$

In (3.7) take $t = W_m$; then (6.5) implies[10]

$$\sum_{(c)} W_m^{p^{nk} - 1} = (-1)^m \frac{D_{m-1}^{p^n - 1}}{D_m} [0, \cdots, m - 1, k], \tag{6.6}$$

and in particular

$$\sum_{(c)} W_m^{p^{nm} - 1} = (-1)^m D_{m-1}^{p^n - 1}. \tag{6.7}$$

For $u_m = 0$, (6.6) yields after some computation

$$\sum_{(c)} U_m^{p^{nk} - 1} = (-1)^m \frac{[0, \cdots, m - 2, k - 1]^{p^n}}{D_{m-1}}, \tag{6.8}$$

[9] For the P-case see II, formula (3.4).

[10] Compare II, p. 943.

while (6.7) gives at once

$$\sum_{(c)} U_m^{p^{nm}-1} = (-1)^m D_{m-1}^{p^n-1}. \tag{6.9}$$

This may also be derived from (6.8).

For another application we turn to (4.1) and replace m by $m + 1$, so that[11]

$$\varphi_{m+1} + C_1\varphi_{m+1}^{p^n} + \cdots + C_k\varphi_{m+1}^{p^{n(k-1)}} = t - \sum_{j=0}^{m} B_{kj}t^{p^{n(k+j)}}, \tag{6.10}$$

and

$$D_m^{p^{nk}} B_{kj} = (-1)^j[0, k, \cdots, k + j - 1, k + j + 1, \cdots, k + m]. \tag{6.11}$$

Now divide both members of (6.10) by $t^{p^{nk}}$ and put $t = W_m = c_0u_0 + \cdots + c_{m-1}u_{m-1} + u_m$. Since the left member vanishes, we get

$$\sum_{(c)} \frac{1}{W_m^{p^{nk}-1}} = \sum_{j=0}^{m} B_{kj}\{\sum_{(c)} W_m^{p^{nj}-1}\}^{p^{nk}}.$$

By (6.6) the only non-vanishing term on the right is the one for which $j = m$; therefore by (6.7) and (6.11),

$$\sum_{(c)} \frac{1}{W^{p^{nk}-1}} = \frac{D_{m-1}^{p^{nk}(p^n-1)}}{D_m^{p^{nk}}} [0, k, \cdots, k + m - 1]. \tag{6.12}$$

Here also we may put $u_m = 0$ and get as the analogue of (6.8):

$$\sum_{(c)}' \frac{1}{U_m^{p^{nk}-1}} = \frac{[0, k + 1, \cdots, k + m - 1]}{D_{m-1}^{p^{nk}}}, \tag{6.13}$$

where in the summation on the left the non-vanishing c_i of greatest subscript is taken $= 1$. In other words the left member

$$= \sum_{i=0}^{m-1} \sum_{(c)} \frac{1}{(c_0u_0 + \cdots + c_{i-1}u_{i-1} + u_i)^{p^{nk}-1}},$$

and the resulting formula may be compared with (6.12).

We remark that by a similar method we may apply (4.7) and (4.9) to derive (6.6) and (6.8).

7. **Some extensions.** As another application of (4.1) we define the inverse of $\varphi_m(t)$. Put

$$\lambda_m(t) = \sum_{i=0}^{\infty} C_i^m t^{p^{ni}}, \tag{7.1}$$

where by (4.2) and (4.4)

$$C_i^m = C_i = \frac{[0, k + 1, \cdots, k + m - 1]}{D_{m-1}^{p^{nk}}}. \tag{7.2}$$

[11] Compare II, p. 945.

From (4.1) we get $\lambda_m(\varphi_m(t)) = t$, and therefore by a general theorem we have also $\varphi_m(\lambda_m(t)) = t$. The question of convergence[12] of (7.1) causes no difficulty and will not be discussed here.

Suppose next that in (2.5) we allow m to become infinite. Put

$$\varphi_m(t) = \sum_{i=0}^{m} (-1)^i B_i^m t^{p^{ni}},$$

so that by (2.4) and (2.3),

$$B_i^m = \frac{[0, \cdots, i-1, i+1, \cdots, m]}{D_{m-1}^{p^n}}.$$

Then formally for

$$\beta_i = \lim_{m=\infty} B_i^m,$$

we get

$$\varphi(t) = \sum_{i=0}^{\infty} (-1)^i \beta_i t^{p^{ni}}. \tag{7.3}$$

Similarly from (7.2) we get

$$\gamma_i = \lim_{m=\infty} C_i^m = \lim_{m=\infty} \frac{[0, k+1, \cdots, k+m-1]}{D_{m-1}^{p^{nk}}},$$

and

$$\lambda(t) = \sum_{i=0}^{\infty} \gamma_i t^{p^{ni}}. \tag{7.4}$$

Then we have

$$\lambda(\varphi(t)) = \varphi(\lambda(t)) = t,$$

so that $\lambda(t)$ is the inverse of $\varphi(t)$.

Also from (2.5) and (7.3) we get

$$\varphi(t) = t \prod_{m=0}^{\infty} \prod_{(c)} \left(1 - \frac{t^{p^n-1}}{(c_0u_0 + \cdots + c_{m-1}u_{m-1} + u_m)^{p^n-1}}\right)$$

or briefly

$$\varphi(t) = t \prod_W{}' \left(1 - \frac{t^{p^n-1}}{W^{p^n-1}}\right), \tag{7.5}$$

where W runs through all linear forms $c_0u_0 + \cdots + c_{m-1}u_{m-1} + u_m$. As an application of the last formula we note

$$\sum_W \frac{1}{W^{p^{nk}-1}} = \gamma_k, \tag{7.6}$$

which may be derived directly from (6.13).

[12] See I, p. 146.

We remark that it does not seem possible to define a quantity having the properties of[13] ξ in the P-case. In that case γ_k of (7.6) satisfies

$$\gamma_k = \frac{\xi^{p^{nk}-1}}{L_k},$$

where L_k is the polynomial (5.11), and more generally

$$\sum \frac{1}{W^m} = B_m \xi^m, \tag{7.7}$$

where B_m is rational. But if we consider the special case

$$u_0 = 1, \qquad u_i = x^{i+1} \qquad \text{for } i > 0,$$

then by direct substitution we may show that (7.7) is false—we find that B_m is *algebraic* but not rational. Since the calculations are rather long, they will be omitted.

8. **The polynomial** $G_k(t)$.[14] For arbitrary $k \geqq 0$ put

$$k = \alpha_0 + \alpha_1 p^n + \cdots + \alpha_s p^{ns} \qquad (0 \leqq \alpha_i < p^n). \tag{8.1}$$

Then we define the polynomial $G_k(t)$ by means of

$$G_k(t) = \prod_{i=0}^{s} f_i^{\alpha_i}(t; u_0, \cdots, u_{i-1}), \tag{8.2}$$

and the closely related $G'_k(t)$ by means of

$$G'_k(t) = \prod_{i=0}^{s} G'_{\alpha_i p^{ni}}(t), \tag{8.3}$$

where

$$G'_{\alpha p^{ni}}(t) = \begin{cases} f_i^{\alpha}(t) & \text{for } 0 \leqq \alpha < p^n - 1, \\ f_i^{\alpha}(t) - f_i^{\alpha}(u_i) & \text{for } \alpha = p^n - 1. \end{cases} \tag{8.4}$$

In particular for $k = p^{nm}$ we have

$$G_{p^{nm}}(t) = f_m(t),$$

while for $k = p^{nm} - 1$ it follows from (2.6) that

$$G'_{p^{nm}-1}(t) = \frac{f_m(t)}{t}, \tag{8.5}$$

and more generally

$$G'_{p^{nm}-p^{ns}}(t) = \frac{f_m(t)}{f_s(t)} \qquad (0 \leqq s \leqq m).$$

[13] I, pp. 150, 161.

[14] For the P-case see this Journal, vol. 6(1940), pp. 486–504. We use the same notation here.

It is readily shown that G_k has the properties

$$(8.6)\qquad \begin{aligned} G_k(ct) &= c^k G_k(t) \qquad\qquad (c \text{ in } GF(p^n)),\\ G_k(t+w) &= \sum_{i+j=k} \binom{k}{i} G_i(t)G_j(w) = (G(t)+G(w))^k, \end{aligned}$$

and that G_k' has the properties

$$(8.7)\qquad \begin{aligned} G_k'(ct) &= c^k G_k(t) \qquad (c \text{ in } GF(p^n),\ c \neq 0),\\ G_k'(t+w) &= (G(t)+G'(w))^k = (G'(t)+G(w))^k. \end{aligned}$$

For $g(t)$ an arbitrary (not necessarily *linear*) polynomial of degree k, consider the representation

$$(8.8)\qquad g(t) = \sum_{i=0}^{k} A_i G_i(t),$$

which obviously exists and is unique. To determine the coefficient A_i we construct the auxiliary polynomial

$$(8.9)\qquad \Phi(t) = \sum_{(c)} g(U_m)\frac{f_m(t)}{t-U_m},$$

where as above

$$U_m = c_0u_0 + \cdots + c_{m-1}u_{m-1}.$$

Clearly the degree of $\Phi(t) = p^{nm}-1$. Also it is easily seen that for arbitrary U_m

$$\Phi(U_m) = (-1)^m D_{m-1}^{p^n-1} g(U_m).$$

Hence we conclude that

$$(8.10)\qquad \Phi(t) = (-1)^m D_{m-1}^{p^n-1} \sum_{i<p^{nm}} A_i G_i(t).$$

Again from (8.5), (8.7) and (8.9) follows

$$\begin{aligned} \Phi(t) &= \sum_{(c)} g(U_m)G'_{p^{nm}-1}(t-U_m)\\ &= \sum_{i+j=p^{nm}-1} G_i(t)\sum_{(c)} g(U_m)G_j'(U_m). \end{aligned}$$

Comparison with (8.10) gives

$$(8.11)\qquad (-1)^m D_m^{p^n-1} A_i = \sum_{(c)} g(U_m)G_{p^{nm}-1-i}(U_m),$$

where $p^{nm} > i$. Thus we have determined the coefficients in (8.8). In much the same way we get also

$$(8.12)\qquad (-1)^m D_m^{p^n-1} A_i = \sum_{(c)} g(W_m)G'_{p^{nm}-1-i}(W_m),$$

where now $p^{nm} > k$ and

$$W_m = c_0u_0 + \cdots + c_{m-1}u_{m-1} + u_m.$$

Parallel to (8.8) we may consider the representation

$$g(t) = \sum_{i=0}^{k} A'_i G'_i(t). \tag{8.13}$$

Using the same method as above we derive the two formulas

$$\begin{aligned}(-1)^m D_m^{p^n-1} A'_i &= \sum_{(c)} g(U_m) G_{p^{nm}-1-i}(U_m) \\ &= \sum_{(c)} g(W_m) G_{p^{nm}-1-i}(W_m),\end{aligned} \tag{8.14}$$

where $p^{nm} > k$. Note that only (8.11) holds for $p^{nm} > i$.

Various other formulas from the P-case are readily extended to the general case. We mention only one or two. For example, we may extend (6.5):

$$\sum_{(c)} \frac{G_k(W_m)}{W_m} = 0 \quad \text{for } k \neq \alpha p^{nm}, 0 \leqq \alpha < p^n,$$

while for $k = \alpha p^{nm}$

$$\sum_{(c)} \frac{G_k(W_m)}{W_m} = f_m^\alpha(u_m) \cdot (-1)^m \frac{D_{m-1}^{p^n}}{D_m} = (-1)^m D_m^{\alpha-1} D_{m-1}^{p^n-\alpha},$$

by (6.3) and (2.6). This result is of use in evaluating

$$\sum_{(c)} \frac{g(W_m)}{W_m}; \tag{8.15}$$

with the notation of (8.8) we find that (8.15)

$$= (-1)^m \sum_{\alpha=0}^{p^n-1} A_{\alpha p^{nm}} D_m^{\alpha-1} D_{m-1}^{p^n-\alpha}.$$

In particular for

$$t^k = \sum_{i=0}^{k} \mathfrak{A}_i^k G_i(t), \tag{8.16}$$

we get for $k \geqq 1$:

$$\sum_{(c)} W_m^{k-1} = (-1)^m D_m^{\lambda-1} D_{m-1}^{p^n-\lambda} \mathfrak{A}_{\lambda p^{nm}}^k,$$

where

$$k \equiv \lambda \pmod{p^n - 1} \qquad (1 \leqq \lambda \leqq p^n - 1).$$

Finally we state:

$$\sum_{(c)} G'_l(U_m) G_k(U_m) = \begin{cases} 0 & \text{for } k + l \neq p^{nm} - 1, \\ (-1)^m D_{m-1}^{p^n-1} & \text{for } k + l = p^{nm} - 1, \end{cases}$$

for $l < p^{nm}$, k arbitrary.

Duke University.

A NOTE ON STRONGLY IRREDUCIBLE MAPS OF AN INTERVAL

By O. G. Harrold, Jr.

A continuous transformation of the compact metric space A onto $B, f(A) = B$, is called *strongly irreducible*[1] provided no proper closed subset of A maps onto all of B. Clearly, any continuous transformation (map) of a compact metric space A into B is strongly irreducible on *some* closed subset A^1 of A. On the other hand, there exists no strongly irreducible map of an *interval* onto a continuum of the form of the letter T. The purpose of this note is to give a proof that a certain extensive class of Peano spaces can be obtained by strongly irreducible maps of the interval. Under the condition given below we not only obtain the desired mapping of the interval I onto the Peano space M but also permit the specification of a certain dense subset of M *in advance* on which the inverse of the mapping function will be single valued.

The letter L will denote throughout the set of local separating points of the Peano continuum M.

THEOREM. *Let M be a Peano continuum such that $M \subset \overline{M - L}$. If P is any countable dense set of non-local separating points of M, there exists a continuous function f such that: $f(I) = M$, where I is the unit interval; if $y \in P$, $f^{-1}(y)$ is a single point; and $\overline{f^{-1}(P)} \supset I$. Thus f is a strongly irreducible map of I onto M.*[2]

Denote by M^I the class of maps of I into subsets of M and suppose it is metrized in the usual way by $\sigma(f,g) = \underset{x \in I}{\text{l.u.b.}}\, \rho[f(x), g(x)]$ $(f, g \in M^I)$. The subset J consisting of maps of I onto all of M constitutes a closed subset of the complete space M^I. By the Hahn-Mazurkiewicz Theorem, $J \neq 0$. Henceforth J will be regarded as our function space. Evidently, J is complete. It will be shown that the class J^* of maps of I onto M which satisfies the conclusions of our theorem is a dense G_δ set in J.

LEMMA 1. *Let M be a Peano continuum such that $M \subset \overline{M - L}$. Let P be any countable dense set of non-local separating points of M. The set $H \subset J$ such that, for $f \in H$, $f^{-1}(y)$ is single valued for each $y \in P$ is a dense G_δ.*

Proof. Set $P = p_1 + p_2 + p_3 + \cdots$, $P_n = p_1 + p_2 + \cdots + p_n$. Denote by F_n^k the subset of J such that for some point of P_n the map has two inverses x^1 and x^2 with $|x^1 - x^2| \geqq 1/k$. The set F_n^k is closed. Put $F_n = \sum_{k=1}^{\infty} F_n^k$. Evidently, F_n is merely the subset of J such that f^{-1} is not single valued on

Received May 22, 1940. The author is a National Research Fellow.

[1] See G. T. Whyburn, *On irreducibility of transformations*, American Journal of Mathematics, vol. 61(1939), p. 820, and references given therein.

[2] Ibid., Theorem 2.

P_n. The set $H_n = J - F_n$ is a G_δ. It will now be shown to be a dense G_δ. Let f be any point of J. Suppose $f^{-1}(p_1) = X$. Let $\epsilon > 0$ be given. Let W be a region in M of diameter $\leqq \epsilon/3 \cdot n$ which contains p_1 and is such that $\overline{W} \cdot (P_n - p_1) = 0$. By continuity, only a finite number of components of $f^{-1}(W)$ can contain points of X. Let these be $U_1, U_2, \cdots, U_k$. On $I - \sum_1^k U_i$, set $g(x) \equiv f(x)$. On each U_i ($i \geqq 2$) define g as a homeomorphism such that (a) $g(U_i) \subset W$, (b) $g(U_i) \cdot p_1 = 0$, and (c) at the end-points of U_i, $f(x) = g(x)$. This is easily done since $M - p_1$ is a region in a Peano space which is uniformly locally connected. Let Y be a Peano subcontinuum in M such that $Y \cdot (P_n - p_1) = 0$, $Y \subset S(\overline{W}, \epsilon/3 \cdot n)$, and $Y \supset \overline{W}$. On U_1 define a continuous transformation g such that (a) $g(\overline{U}_1) = Y$, (b) $g^{-1}(p_1)$ is a single point, and (c) at the end-points of U_1, $f(x) = g(x)$. (See Lemma 2 below.) The constructed function g maps I continuously onto M such that $g^{-1}(p_1)$ is a single point and $\sigma(f, g) < \epsilon/n$. If $g^{-1}(p_2)$ is not single valued, let W be a region in M containing p_2 of diameter $\leqq \epsilon/3 \cdot n$ and such that $\overline{W} \cdot (P_n - p_2) = 0$. If we apply the same procedure, a map g^1 is obtained for which $\sigma(g, g^1) < \epsilon/n$, $(g^1)^{-1}$ is single valued on P_2, and $g^1 \epsilon J$. After at most n steps we arrive at a map $g^0 \epsilon J$ such that each point of P_n has a single inverse and $\sigma(f, g^0) < \epsilon$. Thus H_n is a dense G_δ.

LEMMA 2. *If p is a non-local separating point of the Peano continuum Y and a and b are arbitrary points of $Y - p$, there exists a continuous map of the interval I onto Y such that $f(0) = a$, $f(1) = b$ and $f^{-1}(p)$ is a single point.*

Proof. Let A_n denote the interval $\frac{1}{2} + (n + 3)^{-1} \leqq t \leqq 1$. Since p does not cut Y, we may write $Y - p = \sum_1^\infty B_i$, where B_i is a Peano continuum, $\delta(B_i) \to 0$, $i \to \infty$, and $B_i \cdot B_{i+1} \neq 0$. On A_1 define a continuous function f such that $f(A_1) = B_1$, $f(1) = b$, and $f(\frac{3}{4}) \epsilon B_1 \cdot B_2$. On $A_2 - A_1$ define f so that $f(A_2 - A_1) = B_2$, continuity is preserved at $t = \frac{3}{4}$, and $f(\frac{7}{10}) \epsilon B_2 \cdot B_3$. Continuing in this fashion and setting $f(\frac{1}{2}) = p$, we obtain a map of the interval $\frac{1}{2} \leqq t \leqq 1$ onto Y. By defining, similarly, a map on $0 \leqq t \leqq \frac{1}{2}$ such that $f(0) = a$ and $f(\frac{1}{2}) = p$, we obtain the desired transformation.

LEMMA 3. *Let P be a given countable dense subset of M. The subset D of J such that $f \epsilon D$ implies $\overline{f^{-1}(P)} \supset I$ is a dense G_δ.*

Proof. Let A_i^n be the subset of J for which there is a subinterval U of I of length $\geqq 1/i$ on which there are no inverses to $P_n = p_1 + p_2 + \cdots + p_n$. Denote by B_{ik}^n the subset of A_i^n for which there is an interval U of length $\geqq 1/i$ satisfying the condition $\rho[f(U), P_n] \geqq 1/k$. Then B_{ik}^n is closed and $A_i^n = \sum_{k=1}^\infty B_{ik}^n$. Hence A_i^n is an F_σ. The set $D_i = J - \sum_{n=1}^\infty A_i^n$ is a G_δ consisting entirely of maps of I onto M such that there is no interval U of length $\geqq 1/i$ whose transform does not intersect P. Let $\epsilon > 0$ be given and let f be any point of J. There can be at most k (say) disjoint intervals $U_1, \cdots, U_k$ each

of length $= 1/i$ whose transform under f does not intersect P. A map g^0 of I onto M will be constructed such that $\sigma(f, g^0) < \epsilon$, $g^0 \epsilon D_i$.

Consider the interval U_1. Let the end-points be a and b. Let $\delta > 0$ be such that $| x - y | < \delta$ implies $\rho[f(x), f(y)] < \frac{1}{2}\epsilon$. Let K^1, K^2 and K^3 be the three subintervals (in the order a to b) of U_1 determined by the points $\frac{1}{2}(a + b) - \frac{1}{3}\delta$ and $\frac{1}{2}(a + b) + \frac{1}{3}\delta$. Let S denote a linear transformation sending K^1 into the interval $(a, \frac{1}{2}(a + b))$ with the point a remaining fixed. Let S^1 denote a linear transformation sending K^3 into $(\frac{1}{2}(a + b), b)$ with the point b fixed. Let T denote the continuous transformation resulting by considering these two transformations acting simultaneously. Since P is dense in M, a continuous transformation g can be defined on K^2 such that both end-points of K^2 transform into $f[\frac{1}{2}(a + b)]$, $g(K^2) \cdot P \neq 0$, and $\delta[g(K^2)] < \frac{1}{2}\epsilon$. On the subintervals complementary to K^2 on U_1 set $g(x) \equiv f[T(x)]$. It is easily seen that $g(U_1) \supset f(U_1)$ and if $g(x) = f(x)$ on $I - U_1$, then $\sigma(f, g) < \epsilon$. By similar modifications on $U_2, \cdots, U_k$, a map g^0 is obtained such that $g^0 \epsilon D_i$. Hence the set D_i is a dense G_δ on J. The set $D = \prod_1^\infty D_i$ is comprised wholly of maps such that $\overline{f^{-1}(P)} \supset I$ and is a dense G_δ.

The theorem follows from Lemmas 1 and 3. Hence if M is a Sierpinski triangle curve, a one-dimensional universal curve, or an n-dimensional sphere and P is any given countable dense set of non-local separating points in M, there exists a strongly irreducible mapping of the interval onto M such that the inverse of the mapping function is single valued on P. It is of interest to note that these maps are nowhere arc-preserving[3] in the sense that the only arcs of I which are transformed into arcs are the individual points of I.

THE UNIVERSITY OF VIRGINIA.

[3] G. T. Whyburn, *Arc-preserving transformations*, American Journal of Mathematics, vol. 58(1936), p. 305.

CONJUGATE NETS AND ASSOCIATED QUADRICS

By M. L. MacQueen

1. **Introduction.** This paper is concerned with the study of the projective differential geometry of certain quadrics which are associated with a point of a curve of a conjugate net on an analytic surface in ordinary space. Some portions of the theory of a surface referred to a conjugate net N_x are summarized in §2, where the differential equation of a general conjugate net N_λ, distinct from the parametric conjugate net N_x, is written. The two curves of the net N_λ that pass through a point of the surface will be denoted by C_λ and C_μ. In §3, power series expansions in non-homogeneous projective coördinates for the curve C_λ are computed to terms of as high degree as will be needed in this paper. Some immediate geometrical applications to these series are made. For example, in §4, certain pencils of quadrics having contact of the second order at a point of the surface are investigated. Then, in the following section, the quadrics of Moutard at a point of the surface and in the directions of the tangents to the curves of the net N_λ through the point are considered. Necessary and sufficient conditions for a curve C_λ to be a plane curve and for C_λ to be a cone curve are given in §6. In the next section two quadrics which are associated with each point of a curve C_λ are defined and the equations of these quadrics are found.

2. **Analytic basis.** Let the projective homogeneous coördinates $x^{(1)}, \cdots, x^{(4)}$ of a point P_x on a surface S referred to a conjugate net N_x in ordinary space be given as analytic functions of two independent variables u, v by equations of the form

$$x = x(u, v). \tag{1}$$

Then the coördinates x of a point on the surface and the coördinates y of the point which is the harmonic conjugate of the point x with respect to the foci of the axis of the point x satisfy a system of equations of the form[1]

$$\begin{aligned} x_{uu} &= px + \alpha x_u + Ly, \\ x_{uv} &= cx + ax_u + bx_v, \\ x_{vv} &= qx + \delta x_v + Ny \end{aligned} \qquad (LN \neq 0). \tag{2}$$

It is easily verified that

$$y_u = fx - nx_u + sx_v + Ay, \qquad y_v = gx + tx_u + nx_v + By, \tag{3}$$

Received January 10, 1940.

[1] E. P. Lane, *Projective Differential Geometry of Curves and Surfaces*, Chicago, 1932, p. 138.

where we have placed

$$
(4)\quad
\begin{aligned}
fN &= c_v + ac + bq - c\delta - q_u, & gL &= c_u + bc + ap - c\alpha - p_v,\\
-nN &= a_v + a^2 - a\delta - q, & tL &= a_u + ab + c - \alpha_v,\\
sN &= b_v + ab + c - \delta_u, & nL &= b_u + b^2 - b\alpha - p,\\
A &= b - (\log N)_u, & B &= a - (\log L)_v.
\end{aligned}
$$

Some of the invariants of the parametric conjugate net N_x are given by

$$
(5)\quad
\begin{aligned}
H &= c + ab - a_u, & K &= c + ab - b_v,\\
\mathfrak{H} &= sN, & \mathfrak{K} &= tL,\\
8\mathfrak{B}' &= 4a - 2\delta + l_v, & 8\mathfrak{C}' &= 4b - 2\alpha - l_u,\\
\mathfrak{D} &= -2nL, & r &= N/L,
\end{aligned}
$$

where l is defined by placing

$$(6)\qquad l = \log r.$$

In order that the differential equation

$$(7)\qquad (dv - \lambda du)(dv - \mu du) = 0 \qquad (\lambda = \lambda(u, v),\ \mu = \mu(u, v))$$

may represent a conjugate net on the surface, the two directions defined by this equation must separate harmonically the two asymptotic directions satisfying the equation

$$(8)\qquad L\,du^2 + N\,dv^2 = 0.$$

A necessary and sufficient condition for this is found to be

$$(9)\qquad \mu = -\frac{1}{r\lambda}.$$

The two curves of such a conjugate net N_λ that pass through the point P_x of the surface may be denoted by C_λ and C_μ respectively, according as the direction dv/du has the value λ or μ. We shall suppose that $\lambda \neq 0$ so that the curve C_λ does not belong to the net N_x.

3. **Power series expansions for C_λ.** Let us consider the curve C_λ of the family

$$(10)\qquad dv - \lambda\,du = 0$$

which passes through a point P_x of the surface. Any point X on the curve C_λ and near the point P_x can be defined by the following power series in the increment Δu corresponding to displacement on C_λ from P_x to the point X:

$$(11)\qquad X = x + x'\Delta u + \tfrac{1}{2}x''\Delta u^2 + \tfrac{1}{6}x'''\Delta u^3 + \cdots,$$

where

$$
\begin{aligned}
x' &= x_u + x_v\lambda, \\
x'' &= x_{uu} + 2x_{uv}\lambda + x_{vv}\lambda^2 + x_v\lambda', \qquad (12)\\
x''' &= x_{uuu} + 3x_{uuv}\lambda + 3x_{uvv}\lambda^2 + x_{vvv}\lambda^3 + 3x_{uv}\lambda' + 3x_{vv}\lambda\lambda' + x_v\lambda'' \\
&\qquad (\lambda' = \lambda_u + \lambda\lambda_v).
\end{aligned}
$$

The ray-points x_{-1}, x_1 of the point P_x are given by the formulas

$$x_{-1} = x_u - bx, \qquad x_1 = x_v - ax. \tag{13}$$

If the four points x, x_{-1}, x_1, y are used as the vertices of a local covariant tetrahedron of reference with a suitably chosen unit point, then the point X can be expressed uniquely in the form

$$X = y_1 x + y_2 x_{-1} + y_3 x_1 + y_4 y, \tag{14}$$

where the local coördinates $y_1, \cdots, y_4$ of the point X are represented by the series

$$
\begin{aligned}
y_1 &= 1 + (b + a\lambda)\Delta u + \tfrac{1}{2}A_2\Delta u^2 + \cdots, \\
y_2 &= \Delta u + \tfrac{1}{2}(\alpha + 2a\lambda)\Delta u^2 + \tfrac{1}{6}B_3\Delta u^3 + \cdots, \\
y_3 &= \lambda\Delta u + \tfrac{1}{2}(2b\lambda + \delta\lambda^2 + \lambda')\Delta u^2 + \tfrac{1}{6}C_3\Delta u^3 + \cdots, \\
y_4 &= \tfrac{1}{2}(L + N\lambda^2)\Delta u^2 + \tfrac{1}{6}D_3\Delta u^3 + \tfrac{1}{24}D_4\Delta u^4 + \cdots,
\end{aligned} \tag{15}
$$

and the coefficients $A_2, \cdots, D_4$ are defined by the following formulas:

$$
\begin{aligned}
A_2 &= p + b\alpha + 2(c + 2ab)\lambda + (q + a\delta)\lambda^2 + a\lambda', \\
B_3 &= \alpha_u + \alpha^2 + p - nL + 3\lambda(a_u + c + ab + a\alpha) + 3\lambda^2(a_v + a^2) \\
&\qquad + tN\lambda^3 + 3a\lambda', \\
C_3 &= sL + 3\lambda(b_u + b^2) + 3\lambda^2(b_v + c + ab + b\delta) + \lambda^3(\delta_v + \delta^2 + q + nN) \\
&\qquad + 3\lambda' b + 3\lambda\lambda'\delta + \lambda'', \\
D_3 &= L_u + \alpha L + AL + 3\lambda aL + 3\lambda^2 bN + \lambda^3(N_v + \delta N + BN) + 3\lambda\lambda' N, \qquad (16)\\
D_4 &= L(\alpha_u + \alpha^2 + p - nL) + A(L_u + \alpha L + AL) + (L_u + \alpha L + AL)_u \\
&\quad + 4\lambda[L(a_u + c + ab + a\alpha) + aAL + (aL)_u] \\
&\quad + 6\lambda^2[L(a_v + a^2) + bAN + (bN)_u] \\
&\quad + 4\lambda^3[tLN + A(N_v + \delta N + BN) + (N_v + \delta N + BN)_u] \\
&\quad + \lambda^4[N(\delta_v + \delta^2 + q + nN) + B(N_v + \delta N + BN) + (N_v + \delta N + BN)_v] \\
&\quad + 6aL\lambda' + 12bN\lambda\lambda' + 6\lambda^2\lambda'(N_v + \delta N + BN) + 4N\lambda\lambda'' + 3N\lambda'^2.
\end{aligned}
$$

Introducing non-homogeneous coördinates by the definitions

$$x = y_2/y_1, \qquad y = y_3/y_1, \qquad z = y_4/y_1, \tag{17}$$

we find, by use of the series (17), the following expansions:

$$\begin{aligned} x &= \Delta u + \tfrac{1}{2}(\alpha - 2b)\Delta u^2 + \tfrac{1}{6}[B_3 - 3A_2 - 3(b + a\lambda)(\alpha - 2b)]\Delta u^3 + \cdots, \\ y &= \lambda\Delta u + \tfrac{1}{2}[(\delta - 2a)\lambda^2 + \lambda']\Delta u^2 \\ &\qquad + \tfrac{1}{6}\{C_3 - 3\lambda A_2 - 3(b + a\lambda)[(\delta - 2a)\lambda^2 + \lambda']\}\Delta u^3 + \cdots, \\ z &= \tfrac{1}{2}(L + N\lambda^2)\Delta u^2 + \tfrac{1}{6}[D_3 - 3(b + a\lambda)(L + N\lambda^2)]\Delta u^3 \\ &\qquad + \tfrac{1}{24}[D_4 - 4(b + a\lambda)D_3 - 6(L + N\lambda^2)A_2 \\ &\qquad + 12(b + a\lambda)^2(L + N\lambda^2)]\Delta u^4 + \cdots. \end{aligned} \tag{18}$$

At this point it will be convenient to introduce two functions M, P defined by the following formulas:

$$\begin{aligned} M &= \lambda' + 4\lambda(\mathfrak{C}' - \lambda\mathfrak{B}') + \tfrac{1}{2}\lambda l', \\ P &= M[3r\lambda\lambda' + 4\lambda(r\lambda\mathfrak{C}' + \mathfrak{B}') + \tfrac{1}{2}(r\lambda^2 l' - l' - l_u + r\lambda^3 l_v)] \\ &\qquad - (1 + r\lambda^2)[M' + (1 + r\lambda^2)(\mathfrak{H}/r - \mathfrak{D}\lambda - \mathfrak{K}\lambda^2)]. \end{aligned} \tag{19}$$

Inverting the first of the series (18) to obtain Δu as a power series in x, we find

$$\begin{aligned} \Delta u &= x - \tfrac{1}{2}(\alpha - 2b)x^2 \\ &\qquad - \tfrac{1}{6}[B_3 - 3A_2 - 3(b + a\lambda)(\alpha - 2b) - 3(\alpha - 2b)^2]x^3 + \cdots; \end{aligned} \tag{20}$$

and substituting the result in the last two of the series (18), we arrive at *the following power series expansions for the curve C_λ at the point P_x of the surface*:

$$\begin{aligned} y &= \lambda x + \tfrac{1}{2}Mx^2 + \tfrac{1}{6}a_3x^3 + \cdots, \\ z &= \tfrac{1}{2}(L + N\lambda^2)x^2 + \tfrac{1}{6}Lb_3x^3 + \tfrac{1}{24}Lb_4x^4 + \cdots, \end{aligned} \tag{21}$$

where the coefficients a_3, b_3, b_4 are given by the formulas

$$\begin{aligned} b_3 &= 3r\lambda M + 8(\mathfrak{C}' + r\lambda^2\mathfrak{B}'), \\ a_3 &= \frac{Mb_3 - P}{1 + r\lambda^2}, \\ b_4 &= 4r\lambda a_3 + 3rM^2 + 48r\lambda^2 M\mathfrak{B}' + \frac{24(c_0 + c_1\lambda + c_3\lambda^3 + c_4\lambda^4)}{L}, \end{aligned} \tag{22}$$

and c_0, c_1, c_3, c_4 are defined by placing

$$\begin{aligned} c_0 &= \tfrac{1}{3}L\mathfrak{C}'[12\mathfrak{C}' + (\log \mathfrak{C}'r^{\frac{1}{2}})_u], & c_1 &= \tfrac{1}{6}L(H - \mathfrak{H}), \\ c_4 &= \tfrac{1}{3}N\mathfrak{B}'[12\mathfrak{B}' + (\log \mathfrak{B}'r^{-\frac{1}{2}})_v], & c_3 &= \tfrac{1}{6}N(K - \mathfrak{K}). \end{aligned} \tag{23}$$

4. **Quadrics associated with the curves of the net N_λ.** The equation of any quadric surface having contact of the second order with the surface at the point P_x can be written in the form

$$Lx^2 + Ny^2 - 2z + k_2xz + k_3yz + k_4z^2 = 0, \tag{24}$$

where k_2, k_3, k_4 are arbitrary parameters. By the aid of the power series (21) it is easy to show that the quadric (24) has third-order contact with the curve C_λ at the point P_x in case

$$k_2 + \lambda k_3 = \frac{16(\mathfrak{C}' + r\lambda^3\mathfrak{B}')}{3(1 + r\lambda^2)}. \tag{25}$$

Moreover, the quadric (24) has contact of the fourth order with the curve C_λ at the point P_x if to equation (25) is adjoined the condition

$$Mk_3 + L(1 + r\lambda^2)k_4 = \frac{b_4 - 4r\lambda a_3 - 3rM^2}{3(1 + r\lambda^2)} - \frac{32(\mathfrak{C}' + r\lambda^3\mathfrak{B}')b_3}{9(1 + r\lambda^2)^2}. \tag{26}$$

If k_3 is regarded as arbitrary, we find, using equations (25), (26), that equation (24) can be written in the form

$$\begin{aligned} Lx^2 + Ny^2 - 2z + 16(\mathfrak{C}' + r\lambda^3\mathfrak{B}')xz/3(1 + r\lambda^2) + k_3(y - \lambda x)z \\ + [16M(r^2\lambda^4\mathfrak{B}' - 2r\lambda\mathfrak{C}' + 3r\lambda^2\mathfrak{B}')/3 - 256(\mathfrak{C}' + r\lambda^3\mathfrak{B}')^2/9 \\ + 8(1 + r\lambda^2)(c_0 + c_1\lambda + c_3\lambda^3 + c_4\lambda^4)/L \\ - Mk_3/L(1 + r\lambda^2)]z^2/L(1 + r\lambda^2)^3 = 0. \end{aligned} \tag{27}$$

Therefore *any quadric of the pencil* (27), *where k_3 is arbitrary, has contact of the second order with the surface at the point P_x and contact of the fourth order with the curve C_λ at the point.*

Let us suppose that the quadric (24) which has third-order contact with the curve C_λ at the point P_x also has third-order contact with the curve C_μ at the point. On replacing λ by $-1/r\lambda$ in equation (25), the condition is found to be

$$r\lambda k_2 - k_3 = \frac{16(r^2\lambda^3\mathfrak{C}' - \mathfrak{B}')}{3(1 + r\lambda^2)}. \tag{28}$$

Then by means of equations (25), (28) it is easy to verify the truth of the following statement.

Any quadric of the pencil

$$\begin{aligned} Lx^2 + Ny^2 - 2z + \frac{16(\mathfrak{C}' - \lambda\mathfrak{B}' + r\lambda^3\mathfrak{B}' + r^2\lambda^4\mathfrak{C}')\,xz}{3(1 + r\lambda^2)^2} \\ + \frac{16(\mathfrak{B}' + r\lambda\mathfrak{C}' - r^2\lambda^3\mathfrak{C}' + r^2\lambda^4\mathfrak{B}')\,yz}{3(1 + r\lambda^2)^2} + k_4z^2 = 0, \end{aligned} \tag{29}$$

where k_4 is arbitrary, has contact of the second order with the surface at the point P_x and contact of the third order with both C_λ and C_μ at the point.

There is a unique quadric of the pencil (29) that has fourth-order contact with the curve C_λ at the point P_x. For this quadric the value of k_4 is found by demanding that the series (21) satisfy equation (29) identically in x as far as the terms of the fourth degree. The result is

$$
\begin{aligned}
L(1 + r\lambda^2)^3 k_4 = -8\Big[\tfrac{32}{9}(\mathfrak{C}' + r\lambda^3\mathfrak{B}')^2 + \tfrac{2}{3}M(\mathfrak{B}' + 3r\lambda\mathfrak{C}' \\
- 3r\lambda^2\mathfrak{B}' - r^2\lambda^3\mathfrak{C}') - \frac{(1 + r\lambda^2)(c_0 + c_1\lambda + c_3\lambda^3 + c_4\lambda^4)}{L}\Big].
\end{aligned}
\tag{30}
$$

We shall now find the equation of the Moutard pencil of quadrics[2] in the direction of the tangent to the curve C_λ at the point P_x. A power series expansion[3] for one non-homogeneous coördinate z of a point on the surface in terms of the other two coördinates x, y is given, to terms of the fourth order, by

$$
\begin{aligned}
z = \tfrac{1}{2}(Lx^2 + Ny^2) + \tfrac{4}{3}(L\mathfrak{C}'x^3 + N\mathfrak{B}'y^3) \\
+ c_0x^4 + c_1x^3y + c_3xy^3 + c_4y^4 + \cdots,
\end{aligned}
\tag{31}
$$

where the coefficients c_0, c_1, c_3, c_4 are defined by equations (23). The curve of intersection of the quadric (24) and the surface (31) is known to have a triple point at the point P_x, with triple-point tangents satisfying the equations

$$(\tfrac{4}{3}\mathfrak{C}' - \tfrac{1}{4}k_2)x^3 - \tfrac{1}{4}k_3x^2y - \tfrac{1}{4}rk_2xy^2 + r(\tfrac{4}{3}\mathfrak{B}' - \tfrac{1}{4}k_3)y^3 = 0, \qquad z = 0.$$

If two of the triple-point tangents coincide in the direction of the tangent to the curve C_λ at the point P_x, the third must lie in the direction defined by

$$- \frac{\tfrac{16}{3}\mathfrak{C}' - k_2}{r\lambda^2(\tfrac{16}{3}\mathfrak{B}' - k_3)}, \tag{32}$$

so that k_2, k_3 are connected by the relations

$$
\begin{aligned}
(1 - r\lambda^2)k_2 - 2r\lambda^3k_3 &= \tfrac{16}{3}(\mathfrak{C}' - 2r\lambda^3\mathfrak{B}'), \\
2k_2 + \lambda(1 - r\lambda^2)k_3 &= \tfrac{16}{3}(2\mathfrak{C}' - r\lambda^3\mathfrak{B}').
\end{aligned}
\tag{33}
$$

Solving equations (33) for k_2, k_3 and substituting the results in equation (24) yields *the equation of the Moutard pencil of quadrics in the direction of the tangent to the curve C_λ at the point P_x*,

$$
\begin{aligned}
Lx^2 + Ny^2 - 2z + \frac{16(\mathfrak{C}' + 3r\lambda^2\mathfrak{C}' - 2r\lambda^3\mathfrak{B}')\,xz}{3(1 + r\lambda^2)^2} \\
+ \frac{16(-2r\lambda\mathfrak{C}' + 3r\lambda^2\mathfrak{B}' + r^2\lambda^4\mathfrak{B}')\,yz}{3(1 + r\lambda^2)^2} + k_\lambda z^2 = 0,
\end{aligned}
\tag{34}
$$

where k_λ is arbitrary.

[2] B. Su and A. Ichida, *On certain cones connected with a surface in the affine space*, Japanese Journal of Mathematics, vol. 10(1933), p. 209.

[3] E. P. Lane, *A canonical power series expansion for a surface*, Transactions of the American Mathematical Society, vol. 37(1935), p. 481.

The equations of the triple-point tangent in the direction defined by the formula (32) can be written in the form

$$[(1 - r\lambda^2)\mathfrak{C}' - 2\lambda\mathfrak{B}']x - [(1 - r\lambda^2)\mathfrak{B}' + 2r\lambda\mathfrak{C}']y = 0, \qquad z = 0. \tag{35}$$

If two of the triple-point tangents coincide in the direction of the tangent to the curve C_μ at the point P_x, the third tangent is found to be the same as that defined by equations (35), and the quadrics thus determined form the Moutard pencil of quadrics in the direction μ. The equation of this pencil of quadrics is given by

$$\begin{aligned} Lx^2 + Ny^2 - 2z &+ \frac{16(2\lambda\mathfrak{B}' + 3r\lambda^2\mathfrak{C}' + r^2\lambda^4\mathfrak{C}')\,xz}{3(1 + r\lambda^2)^2} \\ &+ \frac{16(\mathfrak{B}' + 3r\lambda^2\mathfrak{B}' + 2r^2\lambda^3\mathfrak{C}')\,yz}{3(1 + r\lambda^2)^2} + k_\mu z^2 = 0, \end{aligned} \tag{36}$$

where k_μ is arbitrary.

Thus to a tangent at a point P_x of a curve of the net N_λ corresponds a pencil of quadrics such that any quadric of the pencil has second-order contact with the surface at the point and cuts the surface in a curve two of whose triple-point tangents coincide in the direction of the tangent to the curve of the net. Using equations (25) and (34), we find that any quadric of the Moutard pencil in the direction of the tangent to the curve C_λ at the point P_x has third-order contact with the curve C_λ at the point. Moreover, by use of equation (28), we find that any quadric of the pencil (34) also has third-order contact with the curve C_μ at the point P_x in case

$$\mathfrak{B}' + 3r\lambda\mathfrak{C}' - 3r\lambda^2\mathfrak{B}' - r^2\lambda^3\mathfrak{C}' = 0. \tag{37}$$

On replacing λ by dv/du in this equation, we obtain the differential equation of the curves of Segre. Similar results can easily be found for the quadrics of the Moutard pencil (36). Thus the following conclusion is reached:

Any quadric of the Moutard pencil in the direction of the tangent to the curve C_λ (C_μ) at a point P_x of a surface has third-order contact with the curve at the point. It also has third-order contact with the curve C_μ (C_λ) at the point P_x if, and only if, C_λ and C_μ are curves of the Segre-Darboux pencil.

In what follows we shall suppose that C_λ and C_μ are not curves of the Segre-Darboux pencil.

The polar line of the axis at the point P_x of the net N_x with respect to any quadric of the pencil (29) is found to have the equations

$$\begin{aligned} &(\mathfrak{C}' - \lambda\mathfrak{B}' + r\lambda^3\mathfrak{B}' + r^2\lambda^4\mathfrak{C}')x \\ &\qquad + (\mathfrak{B}' + r\lambda\mathfrak{C}' - r^2\lambda^3\mathfrak{C}' + r^2\lambda^4\mathfrak{B}')y = \tfrac{3}{8}(1 + r\lambda^2)^2, \qquad z = 0. \end{aligned} \tag{38}$$

This line crosses the tangent to the curve C_λ at P_x in the point P_λ given by

$$\tfrac{3}{8}(\mathfrak{C}' + r\lambda^3\mathfrak{B}')x + (1 + r\lambda^2)x_{-1} + \lambda(1 + r\lambda^2)x_1, \tag{39}$$

and crosses the tangent to the curve C_μ in the point P_μ defined by

$$\tfrac{3}{8}(\mathfrak{B}' - r^2\lambda^3\mathfrak{C}')x - r\lambda(1 + r\lambda^2)x_{-1} + (1 + r\lambda^2)x_1. \tag{40}$$

The polar line of the axis with respect to a quadric of the Moutard pencil (34) is given by

$$\begin{aligned}(\mathfrak{C}' + 3r\lambda^2\mathfrak{C}' - 2r\lambda^3\mathfrak{B}')x + (-2r\lambda\mathfrak{C}' + 3r\lambda^2\mathfrak{B}' + r^2\lambda^4\mathfrak{B}')y \\ = \tfrac{3}{8}(1 + r\lambda^2)^2, \qquad z = 0.\end{aligned} \tag{41}$$

The two lines defined by equations (38), (41) intersect in the point P_λ. Furthermore, the polar line of the axis with respect to a quadric of the Moutard pencil (36) has the equations

$$\begin{aligned}(2\lambda\mathfrak{B}' + 3r\lambda^2\mathfrak{C}' + r^2\lambda^4\mathfrak{C}')x + (\mathfrak{B}' + 3r\lambda^2\mathfrak{B}' + 2r^2\lambda^3\mathfrak{C}')y \\ = \tfrac{3}{8}(1 + r\lambda^2)^2, \qquad z = 0,\end{aligned} \tag{42}$$

and this line intersects the line (38) in the point P_μ. The two lines defined by equations (41), (42) meet in a point which lies on the triple-point tangent (35).

There is a quadric cone associated with the pencil of quadrics having second-order contact with the surface at the point P_x and third-order contact with both curves of the net N_λ at the point. We shall now define this cone and derive its equation. The polar line of the ray, joining the points x_{-1}, x_1, with respect to any one of the quadrics of the pencil (29) has the equations

$$Lx + k_2z = 0, \qquad Ny + k_3z = 0, \tag{43}$$

where we have placed

$$3(1 + r\lambda^2)^2k_2 = 8(\mathfrak{C}' - \lambda\mathfrak{B}' + r\lambda^3\mathfrak{B}' + r^2\lambda^4\mathfrak{C}'),$$
$$3(1 + r\lambda^2)^2k_3 = 8(\mathfrak{B}' + r\lambda\mathfrak{C}' - r^2\lambda^3\mathfrak{C}' + r^2\lambda^4\mathfrak{B}').$$

When λ varies, this line generates a quadric cone with its vertex at the point P_x. The equation of this cone is found, by eliminating λ from equations (43), to be

$$Lx^2 + Ny^2 + 4\mathfrak{C}'xz + 4\mathfrak{B}'yz + \frac{32(\mathfrak{B}'^2 + r\mathfrak{C}'^2)\,z^2}{9N} = 0. \tag{44}$$

Since $\mathfrak{B}'^2 + r\mathfrak{C}'^2 \neq 0$, so that the surface is not ruled, the discriminant of this cone does not vanish. Thus we prove the following theorem:

At a point P_x of a surface the polar line of the ray of the net N_x with respect to any quadric of the pencil (29) *generates the quadric cone* (44) *when λ varies.*

It is evident that this cone is intersected by the tangent plane at the point P_x of the surface in the asymptotic tangents through the point. Moreover, the two planes tangent to the cone along the asymptotic tangents are found to intersect in the associate axis

$$Lx + 2\mathfrak{C}'z = 0, \qquad Ny + 2\mathfrak{B}'z = 0.$$

We may regard a line l_1 through a point P_x of the surface as a line which joins the point P_x to the point

$$-dx_{-1} - ex_1 + y,$$

wherein d, e are functions of u, v. Such a line l_1 has been defined[4] to be a canonical line of the first kind in case

$$dL = h\mathfrak{C}', \qquad eN = h\mathfrak{B}',$$

where h is a constant. A canonical line of the first kind lies in the canonical plane whose equation is

$$\mathfrak{B}'x - r\mathfrak{C}'y = 0. \tag{45}$$

A canonical line of the second kind crosses the parametric tangents at the point P_x in the points

$$x_{-1} - h\mathfrak{C}'x, \qquad x_1 - h\mathfrak{B}'x.$$

Thus the two lines of intersection of the cone (44) and the canonical plane (45) are lines of the first canonical pencil, and the values of h which characterize these lines are found by solving equations (44), (45) simultaneously. It is now easy to verify the conclusion:

The cone (44) *is intersected by the canonical plane in two canonical lines of the first kind for which* $h = \frac{4}{3}$ *and* $h = \frac{8}{3}$.

The line for which $h = \frac{4}{3}$ seems to be new. Thus another characterization of the line for which $h = \frac{8}{3}$ is found, which should be compared with that given[5] by Lane and MacQueen. Finally, the planes tangent to the cone along these two canonical generators are found to intersect in the line

$$\mathfrak{C}'x + \mathfrak{B}'y = 0, \qquad z = 0,$$

which has been called[6] the second canonical tangent of the conjugate net and its associate conjugate net. Hence we have the theorem:

The planes tangent to the cone (44) *along its lines of intersection with the canonical plane intersect in the second canonical tangent.*

We shall next consider a cubic curve in the tangent plane at the point P_x of the surface. The homogeneous local coördinates of the point P_λ given by the formula (39) are

$$y_1 = \tfrac{8}{3}(\mathfrak{C}' + r\lambda^3\mathfrak{B}'), \quad y_2 = 1 + r\lambda^2, \quad y_3 = \lambda(1 + r\lambda^2), \quad y_4 = 0. \tag{46}$$

When λ varies, the locus of this point is a nodal cubic curve whose equations are found, by eliminating λ homogeneously from equations (46), to be

$$x^2 + ry^2 - \tfrac{8}{3}(\mathfrak{C}'x^3 + r\mathfrak{B}'y^3) = 0, \qquad z = 0. \tag{47}$$

[4] W. M. Davis, *Contributions to the theory of conjugate nets*, Chicago doctoral dissertation (1932), p. 18.

[5] E. P. Lane and M. L. MacQueen, *The curves of a conjugate net*, this Journal, vol. 5(1939), p. 699.

[6] W. M. Davis, loc. cit., p. 19.

The locus of the point P_μ, defined by the formula (40), is found to be the same curve. Several interesting results can be easily established. This cubic has a double point at P_x with the asymptotic tangents at P_x for double-point tangents. The cubic has three points of inflexion which lie on the associate ray

$$y_1 - 2\mathfrak{C}'y_2 - 2\mathfrak{B}'y_3 = 0, \qquad y_4 = 0,$$

some one of the inflexions lying on each of the tangents of Darboux at P_x. The tangents to this cubic at the points P_λ, P_μ are respectively the lines defined by equations (41), (42). The points of intersection, distinct from the point P_x, of the cubic with the parametric tangents through P_x are the points

$$x_{-1} + \tfrac{8}{3}\mathfrak{C}'x, \qquad x_1 + \tfrac{8}{3}\mathfrak{B}'x,$$

which determine the principal join of the parametric curves at P_x. Moreover, the tangents to the cubic at these points pass respectively through the ray points x_1, x_{-1} and meet in the line

$$\mathfrak{C}'x - \mathfrak{B}'y = 0, \qquad z = 0,$$

which is the harmonic conjugate of the second canonical tangent with respect to the parametric tangents at the point P_x. Finally, the cubic crosses the associate conjugate tangents at the point P_x in points, other than P_x, defined by the formulas

$$x_{-1} + \tfrac{4}{3}\mathfrak{C}'x \pm (x_1 + \tfrac{4}{3}\mathfrak{B}'x)r^{-\frac{1}{2}}.$$

The line which passes through these two points is a canonical line of the second kind for which $h = -\frac{4}{3}$.

5. **The quadrics of Moutard for the tangents to the curves C_λ, C_μ.** The quadric of Moutard at a point P_x of the surface and in the direction of the tangent to the curve C_λ through P_x is defined to be the locus of the osculating conics at the point P_x of the curves of section of the surface made by planes through this tangent. In order to obtain the equation of this quadric we proceed in the usual manner. The equation

$$z = k(y - \lambda x) \qquad (k \neq 0) \tag{48}$$

represents an arbitrary plane through the tangent to the curve C_λ at the point P_x. This plane cuts the surface (31) in a plane curve whose projection from the point P_y onto the tangent plane, $z = 0$, of the surface at the point P_x is a curve C. Without recording the details of the calculation, it is sufficient to state that the equation, in the tangent plane, of the curve C is found by eliminating z between equations (31), (48), and then solving for y as a power series in x. On calculating the equation of the osculating conic of the curve C at the point P_x and eliminating k between this result and equation (48), we arrive at

the equation of the quadric of Moutard for the tangent to the curve C_λ at the point P_x, namely,

$$
\begin{aligned}
(49)\qquad & Lx^2 + Ny^2 - 2z + \frac{16(\mathfrak{C}' + 3r\lambda^2\mathfrak{C}' - 2r\lambda^3\mathfrak{B}')\,xz}{3(1 + r\lambda^2)^2} \\
&\qquad + \frac{16(-2r\lambda\mathfrak{C}' + 3r\lambda^2\mathfrak{B}' + r^2\lambda^4\mathfrak{B}')\,yz}{3(1 + r\lambda^2)^2} + k_\lambda z^2 = 0,
\end{aligned}
$$

where k_λ is given by

$$
(50)\quad L(1 + r\lambda^2)^3 k_\lambda = -8\left[\tfrac{32}{9}(\mathfrak{C}' + r\lambda^3\mathfrak{B}')^2 - \frac{(1 + r\lambda^2)(c_0 + c_1\lambda + c_3\lambda^3 + c_4\lambda^4)}{L}\right].
$$

The equation of the quadric of Moutard for the tangent to the curve C_μ at the point P_x can be written by replacing λ by $-1/r\lambda$ in equation (49). The result is found to be

$$
\begin{aligned}
(51)\qquad & Lx^2 + Ny^2 - 2z + \frac{16(2\lambda\mathfrak{B}' + 3r\lambda^2\mathfrak{C}' + r^2\lambda^4\mathfrak{C}')\,xz}{3(1 + r\lambda^2)^2} \\
&\qquad + \frac{16(\mathfrak{B}' + 3r\lambda^2\mathfrak{B}' + 2r^2\lambda^3\mathfrak{C}')\,yz}{3(1 + r\lambda^2)^2} + k_\mu z^2 = 0,
\end{aligned}
$$

where the value of k_μ is given by

$$
\begin{aligned}
(52)\qquad & L(1 + r\lambda^2)^3 k_\mu \\
&\quad = -8\left[\frac{32(\mathfrak{B}' - r\lambda^3\mathfrak{C}')^2}{9r} - \frac{(1 + r\lambda^2)(c_0 r^4\lambda^4 - c_1 r^3\lambda^3 - c_3 r\lambda + c_4)}{Lr^2}\right].
\end{aligned}
$$

The quadrics of Moutard (49), (51) are intersected by the tangent plane at the point P_x of the surface in the asymptotic tangents at the point. Furthermore, the cone projecting the curve of intersection of these two quadrics from the point P_x consists of two planes, one of which is the tangent plane of the surface at the point. The other plane, which contains the conic of intersection of the two quadrics, has the equation

$$
\begin{aligned}
(53)\qquad & (\mathfrak{C}' - 2\lambda\mathfrak{B}' - r\lambda^2\mathfrak{C}')\,x - (\mathfrak{B}' + 2r\lambda\mathfrak{C}' - r\lambda^2\mathfrak{B}')\,y \\
&\qquad + \frac{3(1 + r\lambda^2)(k_\lambda - k_\mu)\,z}{16} = 0,
\end{aligned}
$$

where k_λ, k_μ are defined by equations (50), (52). It follows that the plane which contains the conic of intersection of the quadrics of Moutard for the tangents to the curves C_λ and C_μ at a point of the surface intersects the tangent plane of the surface at the point in the triple-point tangent (35).

Incidentally, when λ varies, it is known[7] that the plane (53) envelops a cone of the third class whose cusp planes pass through the Segre tangents at the point P_x and intersect in the first directrix of Wilczynski.

[7] Louis Green, *Systems of quadrics associated with a point of a surface*, American Journal of Mathematics, vol. 60(1938), p. 661.

If the parameter k_3 in equation (27) has the value

$$k_3 = \frac{16(-2r\lambda\mathfrak{C}' + 3r\lambda^2\mathfrak{B}' + r^2\lambda^4\mathfrak{B}')}{3(1 + r\lambda^2)^2}, \tag{54}$$

the resulting equation is found to be the same as the equation of the quadric of Moutard (49). Therefore, *the quadric of Moutard for the tangent to the curve C_λ at the point P_x is a unique quadric of the pencil* (27) *and for this quadric the parameter k_3 has the value given by equation* (54). Moreover, if we demand that equation (34) be satisfied by the series (21) identically in x up to, and including, the terms of the fourth degree, we obtain for k_λ the value given in equation (50). Consequently *the quadric of Moutard* (49) *is the unique quadric of the Moutard pencil* (34) *that has fourth-order contact with the curve C_λ at the point P_x*. Finally, the quadric (29) that has fourth-order contact with the curve C_λ at the point P_x meets the quadric of Moutard (49) in the asymptotic tangents through P_x and in a conic which lies in the osculating plane of the curve C_λ at P_x, the equation of this plane being given by

$$(L + N\lambda^2)(y - \lambda x) - Mz = 0. \tag{55}$$

6. **Conditions for plane curves and cone curves.** The curve C_λ is a plane curve if, at every point of the curve, the osculating plane at the point has third-order contact with the curve. Demanding that the series (21) satisfy equation (55) identically in x up to, and including, the term in x^3, we find that *the curve C_λ is a plane curve if, and only if, $P = 0$.* If λ is replaced by dv/du in this condition, the result is an equation of the third order for v as a function of u along a plane curve C_λ.

The osculating plane at every point of a curve C_λ contains the axis of the net N_x at the point if, and only if, $M = 0$, so that this is the differential equation of the union curves of the axis congruence. If $M = P = 0$, inspection of equations (19) leads to the well-known result that a union curve of the axis congruence is a plane curve if, and only if, it is an axis curve.

Dual considerations lead to cone curves which are defined to be the curves of contact of cones circumscribed about the surface. The tangent planes at three consecutive points of the curve C_λ intersect in the corresponding ray-point of C_λ given by

$$-\frac{Wx}{\lambda} + (1 + r\lambda^2)x_{-1} + \mu(1 + r\lambda^2)x_1, \tag{56}$$

where W is defined by

$$W = \lambda' - 4\lambda(\mathfrak{C}' - \lambda\mathfrak{B}') + \tfrac{1}{2}\lambda l'. \tag{57}$$

The tangent planes at every point of the curve C_λ envelop a developable surface whose edge of regression is the locus of the ray-point of the curve C_λ. If this developable is a cone, then the ray-point is fixed as the point P_x varies along the curve C_λ, so that the corresponding total derivative with respect to u of

the expression (56) is proportional to the expression itself. From this proportionality we find that *the curve C_λ is a cone curve in case*

$$
\begin{aligned}
(58)\quad & W[3r\lambda\lambda' - 4\lambda(r\lambda\mathfrak{C}' + \mathfrak{B}') + \tfrac{1}{2}(r\lambda^2 l' - l' - l_u + r\lambda^3 l_v)] \\
& \qquad - (1 + r\lambda^2)[W' + (1 + r\lambda^2)(H/r - \lambda\mathfrak{D} - \lambda^2 K)] = 0.
\end{aligned}
$$

7. Quadrics associated with a one-parameter family of curves. At a point P_x of the u-curve and at two neighboring points P_1, P_2 on this curve, let us construct the tangents of each of the curves of the family

$$(59)\qquad dv - \lambda\, du = 0$$

which pass through the points. These three tangents determine a quadric surface. The limit of this quadric, as P_1, P_2 independently approach P_x along the u-curve, is a quadric which will be denoted by Q_u. A second quadric, Q_v, can be defined similarly by using the tangents to the curves of the family (59) at three consecutive points of the v-curve through the point P_x.

It may be remarked that D. Sun has defined[8] in a similar way two quadrics which are associated with a one-parameter family of curves on a surface. However, in the definition just stated, he uses the two asymptotic curves at a point of the surface instead of the two curves of a given conjugate net.

The coördinates X of any point P_1 on the u-curve and near the point P_x are given by an expansion of the form

$$(60)\qquad X = x + x_u \Delta u + \tfrac{1}{2} x_{uu} \Delta u^2 + \cdots .$$

The equation of the curve of the family (59) through the point P_1 is given by

$$(61)\qquad dv - \lambda(u + \Delta u, v)\, du = 0,$$

where

$$(62)\qquad \lambda(u + \Delta u, v) = \lambda + \lambda_u \Delta u + \tfrac{1}{2}\lambda_{uu} \Delta u^2 + \cdots .$$

The coördinates of a point Y on the tangent to the curve (59) at the point P_1 are given by

$$
\begin{aligned}
(63)\quad & Y = x_u + x_v\lambda(u + \Delta u, v) + [x_{uu} + x_{uv}\lambda(u + \Delta u, v)]\Delta u \\
& \qquad + \tfrac{1}{2}[x_{uuu} + x_{uuv}\lambda(u + \Delta u, v)]\Delta u^2 + \cdots .
\end{aligned}
$$

Any point Z on this tangent is defined by a linear combination of the form

$$(64)\qquad Z = hX + kY \qquad (h, k \text{ scalars}).$$

If the points x, x_u, x_v, y are used as the vertices of a local tetrahedron of reference with a suitably chosen unit point, we find that the local coördinates

[8] D. Sun, *Parametric osculating quadrics of a one-parameter family of curves on a surface*, Tôhoku Mathematical Journal, vol. 32(1930), p. 81.

$\zeta_1, \cdots, \zeta_4$ of the point Z are given, to terms of as high degree in Δu as will be needed, by the series

$$
\begin{aligned}
\zeta_1 &= h + k(p + c\lambda)\Delta u + \tfrac{1}{2}[hp + 2kc\lambda_u + k(p_u + \alpha p + fL) \\
&\qquad + k(c_u + bc + ap)\lambda]\Delta u^2 + \cdots, \\
\zeta_2 &= k + [h + k(\alpha + a\lambda)]\Delta u + \tfrac{1}{2}[h\alpha + 2ka\lambda_u + k(\alpha_u + \alpha^2 + p - nL) \\
&\qquad + k(a_u + c + ab + a\alpha)\lambda]\Delta u^2 + \cdots, \\
\zeta_3 &= k\lambda + k(b\lambda + \lambda_u)\Delta u + \tfrac{1}{2}k[sL + (b_u + b^2)\lambda + 2b\lambda_u + \lambda_{uu}]\Delta u^2 + \cdots, \\
\zeta_4 &= kL\Delta u + \tfrac{1}{2}[hL + k(L_u + \alpha L + AL) + kaL\lambda]\Delta u^2 + \cdots.
\end{aligned} \tag{65}
$$

Demanding that the equation of a general quadric be satisfied by the series (65) for $\zeta_1, \cdots, \zeta_4$ identically in h, k and identically in Δu as far as the terms of the second degree, we obtain the equation of the quadric Q_u referred to the tetrahedron x, x_u, x_v, y, namely,

$$
\begin{aligned}
&L\lambda^2 y_2^2 - Ly_3^2 - 2\lambda^2 y_1 y_4 - \lambda^2[\alpha + (\log r\lambda)_u]y_2 y_4 \\
&\qquad + \lambda[4\mathfrak{C}' + 3(\log \lambda r^{\frac{1}{2}})_u - 2a\lambda]y_3 y_4 + \lambda^2 D y_4^2 = 0,
\end{aligned} \tag{66}
$$

the coefficient D being defined by

$$
LD = 4\mathfrak{C}'_u + (\log \lambda r^{\frac{1}{2}})_{uu} - (\log \lambda)_u[4\mathfrak{C}' + (\log \lambda r^{\frac{1}{2}})_u] + H\lambda + \frac{\mathfrak{H}}{r\lambda}. \tag{67}
$$

In order to write the equation of the quadric Q_u referred to the tetrahedron x, x_{-1}, x_1, y it is sufficient to replace y_1 in equation (66) by $y_1 - by_2 - ay_3$. Thus we find that *the equation of the quadric Q_u at the point P_x, referred to the tetrahedron x, x_{-1}, x_1, y, is*

$$
\begin{aligned}
&L\lambda^2 y_2^2 - Ly_3^2 - 2\lambda^2 y_1 y_4 + \lambda^2[J - 2(\log \lambda r^{\frac{1}{2}})_u]y_2 y_4 \\
&\qquad + \lambda[J + 2(\log \lambda r^{\frac{1}{2}})_u]y_3 y_4 + \lambda^2\left[J_u - (\log \lambda)_u J + H\lambda + \frac{\mathfrak{H}}{r\lambda}\right]\frac{y_4^2}{L} = 0,
\end{aligned} \tag{68}
$$

where J is defined by placing

$$
J = 4\mathfrak{C}' + (\log \lambda r^{\frac{1}{2}})_u. \tag{69}
$$

The equation of the quadric Q_v at the point P_x can be written by making the appropriate symmetrical interchanges of the symbols. The result is

$$
\begin{aligned}
&N\lambda^2 y_2^2 - Ny_3^2 + 2y_1 y_4 - \lambda[I - 2(\log \lambda r^{\frac{1}{2}})_v]y_2 y_4 \\
&\qquad - [I + 2(\log \lambda r^{\frac{1}{2}})_v]y_3 y_4 - \left[I_v + (\log \lambda)_v I + \frac{K}{\lambda} + r\lambda\mathfrak{K}\right]\frac{y_4^2}{N} = 0,
\end{aligned} \tag{70}
$$

where I is given by

$$
I = 4\mathfrak{B}' - (\log \lambda r^{\frac{1}{2}})_v. \tag{71}
$$

We shall next prove a theorem concerning the polar relation with respect to the quadrics Q_u, Q_v thus associated with a point P_x of a curve C_λ of the family

(59). As u, v vary, an arbitrary line l_1 through the point P_x and the point $(0, -d, -e, 1)$ generates a congruence Γ_1. The reciprocal polar lines of this line l_1 with respect to the quadrics Q_u, Q_v are respectively

$$(72)\quad \begin{aligned} &2\lambda^2 y_1 - \lambda^2[J - 2\,dL - 2(\log \lambda r^{\frac{1}{2}})_u]y_2 \\ &\qquad - [\lambda J + 2eL + 2\lambda(\log \lambda r^{\frac{1}{2}})_u]y_3 = 0, \; y_4 = 0, \\ &2y_1 - [\lambda I + 2\lambda^2\,dN - 2\lambda(\log \lambda r^{\frac{1}{2}})_v]y_2 \\ &\qquad - [I - 2eN + 2(\log \lambda r^{\frac{1}{2}})_v]y_3 = 0, \; y_4 = 0. \end{aligned}$$

If we demand that these two lines shall coincide we find, after some simplification, that λ satisfies the equations

$$(73)\quad \begin{aligned} &2d(L + N\lambda^2) - 4(\mathfrak{C}' - \lambda\mathfrak{B}') + (\log \lambda r^{\frac{1}{2}})_u - 3\lambda(\log \lambda r^{\frac{1}{2}})_v = 0, \\ &2e(L + N\lambda^2) + 4\lambda(\mathfrak{C}' - \lambda\mathfrak{B}') + 3\lambda(\log \lambda r^{\frac{1}{2}})_u - \lambda^2(\log \lambda r^{\frac{1}{2}})_v = 0. \end{aligned}$$

Multiplying the first of equations (73) by $-\lambda$ and adding the resulting equation to the second of equations (73), we obtain a result which may be written in the form

$$(74)\quad \lambda' = -eL - (4\mathfrak{C}' + \tfrac{1}{2}l_u - dL)\lambda + (4\mathfrak{B}' - \tfrac{1}{2}l_v - eN)\lambda^2 + dN\lambda^3.$$

When λ is replaced by dv/du, this equation becomes an equation of the second order for v as a function of u along a union curve of the congruence Γ_1. So we have proved the theorem:

If a curve C_λ of the family (59) *is such that at each of its points an arbitrary line l_1 of a congruence Γ_1 has the same polar line with respect to the quadric Q_u as with respect to the quadric Q_v, then the curve C_λ is a union curve of the congruence Γ_1.*

SOUTHWESTERN COLLEGE.